M. I. Kaganov

Electrons

Phonons

$a \approx 10^{-8}$ cm
$e \approx 4.8 \times 10^{-10}$ CGSE
$\varepsilon(\mathbf{p} + 2\pi\hbar\mathbf{b}) = \varepsilon(\mathbf{p})$
$\tfrac{1}{2}\hbar$;
$\varepsilon = \hbar\omega$; $R = \frac{1}{nec}$ $k = \frac{2\pi}{\lambda}$, $\Omega \approx \frac{u\pi\hbar}{a}$
$\mathbf{p} = \hbar\mathbf{k}$; $\omega = sk$; $\hbar \approx 10^{-27}$ erg·sec
$s \approx 10^5$ cm/sec; $p_F = \hbar/a = v_F m_e$
$\varepsilon_F = (3\pi^2 n_e)^{1/3} \frac{\hbar^2}{2m_e}$; $m_e \approx 9.1 \times 10^{-28}$
$\omega = gH \leftarrow$
ε
$\sigma = ne^2 l/p_F = ne^2 u$
$c = 3 \times 10^{10}$ cm/sec
$\Delta x \Delta p_x \gtrsim \tfrac{1}{2}\hbar$

ε
$\Delta x \Delta p_x$
$\hbar$;
$p_2 \varepsilon_2$
$u_p = 1840 m_e$
$n_e \approx z/a^3$
$\mu = \frac{e\hbar}{2mc}$
$ak \ll 1$
$\varepsilon\omega$
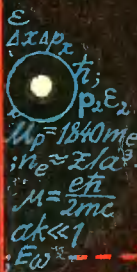
Magnons

Mir
Publishers
Moscow

М. И. Каганов

Электроны. Фононы. Магноны

M. I. Kaganov

Electrons

Phonons

Magnons

Translated from
the Russian
by V. I. Kissin

Contents

Instead of an Introduction

Languages of Science

When science perceives the surrounding world and transforms "things in themselves into things for us", when it masters new fields and turns its achievements into everyday tools of humanity, it also fulfills one additional function. Namely, it *composes a picture of the world* which is modified by each subsequent generation and serves as one of the most important characteristics of civilization. The picture of the world, that is, the sum total of humanity's information about nature, is stored in hundreds of volumes of special monographs and in tens of thousands of articles in scientific journals.

Strictly speaking, this picture is known to humanity as a whole but not to any single person. One man, even with the best possible education, knows the details of only a small fragment of the total picture and is satisfied with approximate information about everything that lies beyond his special field.

The difficulties in obtaining the picture of the surrounding world result not only from the unlimited diversity of data but also from the

existence of specialized languages. These languages are means of communicating and developing logical structures within separate scientific domains; these languages are totally devoid of meaning for a scientist working in a remote field, and are only approximately comprehensible to those working in an adjacent field. Generalization of scientific results and the composition of the picture of the world demand that the descriptions be translated from a specialized language into an ordinary (universal) language. And it is here that the obstacle is encountered: specialized languages are formalized to a much higher extent than any ordinary everyday language.

The translation is always difficult. This is especially true for a translation from a scientific language into a language in which the meaning of each concept is not strictly defined, but can be easily modified under the influence of the experience accumulated by the user of the language. Pictures of the world created in the minds of different people are different not only because the people have digested unequal amounts of information but also because this information is encoded into different languages. A biologist constructs a world picture quite different from that of a physicist. An engineer's picture is much more "mechanistic" than that of a specialist in the humanities.

Popular science literature is an attempt to translate from a rigorous scientific language into a less formal language.

It would be wrong to think that scientists, speaking or thinking about their professional topics, always employ only rigorously formalized

scientific terminology. Far from it. It is not difficult to find, by listening to arguments in a scientific discussion, by paying attention to the wording of reports delivered at conferences and seminars, and by simply listening to specialists talk informally during these conferences, that each branch of science generates two languages. One rigorous and precise, the other much less rigorous.

This second language is a mixture of special terms and everyday words. Repeated use of the latter gives to these words a very peculiar meaning that is unlikely to be found even in the best encyclopaedic dictionary. What is essential, however, is that the addition of the scientific meaning does not suppress the emotive flavour of the word.

There can be no doubt that the words of any human language possess a magic power of stirring up chains of associative images, prodding the mind, and stimulating the emotions. This makes the word a very powerful tool. This explains why a scientist searching for a strict solution uses lively colloquialisms when arguing with the opponents of his point of view; he does not restrict his language to a scientific lingo whose words are precise but lack emotion.

Popular science literature familiarizes the reader with the "colloquial" language of science.

Conventional words often entail allusions which are marginal in a scientific context and therefore interfere with the understanding of a statement. The translation from the scientific language to a conventional language gives rise to losses. Precision is sacrificed, which is an unavoidable price of simplification. It is pos-

sible, however, to try and reduce the *load of unnecessary allusions* that trail after each everyday word.

Take the verb "to decay". A nonphysicist learns from the unabridged edition of the *Random House Dictionary of the English Language* (1966) that "to decay" means:

> vi. 1. to decline in excellence, prosperity, health, etc.; deteriorate. 2. to become decomposed; rot. 3. (*phys.*) (of a radioactive nucleus) to change spontaneously into one or more different nuclei in a process in which particles, as alpha particles, are emitted from the nucleus, electrons are captured or lost, or fission takes place.

A physicist will try to explain, however, that the *decay* of a neutron into a proton, electron, and antineutrino does not mean that prior to the decay the neutron was composed (in separate parts) of the proton, electron, and antineutrino. The word "decay" means here, the physicist says, the "transformation", despite everybody referring to it as "decay".

Another example: "collision". The same dictionary states:

> n. 1. the act of colliding; a coming violently into contact; crash (as railway trains or ships). 2. a clash; conflict. 3. (*phys.*) the meeting of particles or bodies in which each exerts a force upon the other, causing the exchange of energy or momentum.

But in solid state physics an electron-phonon *collision* means that the electron has "absorbed" the phonon.

A comical analogy: a wolf-hare collision. After the collision, the wolf is alone in the field.

Science gives rise to new concepts almost everyday, so that new terms have to be created constantly. Words from ordinary everyday language are often borrowed to produce these terms. It is popular nowadays to borrow words from a domain far removed from science. The physics of elementary particles, for example, has incorporated "strangeness", "charm", "colour", and "flavour".

This vogue may be connected not so much with the unrestrained inventiveness of the creators of new physics as with their attempts to avoid the concomitant introduction of unwanted concepts. The author of the name "quark" for a subnucleon particle (M. Gell-Mann) was very conscious (or felt intuitively) that the allusions surrounding Joyce's quarks cannot affect the notion of the properties of the (then) hypothetical subparticle.

Popular science literature helps the layman to perceive the scientific content of words that were extracted from everyday language and transferred to unfamiliar surroundings.

But the main target of popular science literature is, of course, to acquaint a great number of readers with the progress of science.

The book you are reading now is a popular science book on the quantum physics of the solid state. We are aware of the numerous popular science books devoted to solid state physics. The present book is different in that it is an attempt to concentrate *exclusively* on the quantum physics of the solid state and to disregard applications. It is a book about the methods of interpreting macroscopic effects, about the relationship between solid state physics and quan-

tum mechanics, about the creation and use of new concepts ... . I have also pursued another goal—to "lift the veil" by explaining *how* the results are obtained and not restricting the presentation to an itemization of the results. The reader will be the judge of whether the attempt has failed or succeeded.

Writing this book was a pleasure, and I "tortured" my friends and relations, and above all my wife, by having them listen to my reading some of the passages aloud. They were always patient listeners, and deserve my deepest gratitude.

Chapter 1

On Physics in General and Quantum Mechanics in Particular

To understand means to simplify

Man comprehends the world in the process of his never-ending adaptation of his surroundings to his needs. The results of this activity are both the things the man produces and a set of ideas about the surrounding world. On the one hand, this set of notions becomes continuously more and more detailed (this is a result of our studying the world), and, on the other hand, the concept of this world becomes more and more generalized. Let us discuss the latter aspect of the process of cognition.

People usually think that with each new theory and the ensuing formulation of new concepts, the picture of the world grows more complicated. It is often said that the physics of today is fantastically complicated while it was much simpler in the past. This attitude was even expressed in a quatrain consisting of an eighteenth century epigram:

Nature, and nature's laws, lay hid in night:
God said, *Let Newton be*! and all was light.
(*Alexander Pope*)

followed by an epigram written in our time:

It did not last: the Devil, howling *Ho!*
Let Einstein be! restored the status quo.
<div align="right">(J. C. Squire)</div>

It is indeed true that the theories of modern physics resort to complicated (and sometimes extremely complicated) mathematical techniques. Furthermore, physics becomes less and less visually clear and more and more removed from everyday experience. In order to discover a new phenomenon, a theoretical physicist now needs sophisticated mathematical tools, while experimenters need either giant accelerators, or staggeringly strong magnetic fields, or extremely low temperatures, or even all three together. Conclusions from physical theories sometimes look more fantastic than anything a science fiction writer may invent. (Isn't the dependence of time flow on the choice of the frame of reference a sufficiently conclusive example?) But nevertheless, all these complications essentially lead to a simpler picture of the world.

The concept of the world, moulded by thousands of run-of-the-mill scientists and by a few geniuses, is, in its ultimate form (or in the form considered ultimate today), an impressively elegant and simple structure: thousands of phenomena which at first glance appear unrelated are shown to be logical corollaries (and hence, mathematically derived corollaries) of a limited number of statements. An explanation is found for the diversity itself (and therefore for the complexities of the picture as a whole). And as a rule, no fundamental ideas need to be added to explain the observed variety of phenomena.

This variety is a result of the same fundamental ideas.

J. Kepler found his famous simple laws by observing the motion of planets in the solar system. The laws are simple because all the diversity in the motions of the planets was reduced to three arithmetical relations. But this led to a new complication, this time of an "ideological" character. What was the origin of these relations: the divine wisdom of nature or something more finite? We know that the answer was given by Isaac Newton when he developed Newtonian mechanics and formulated the law of gravitation. No simplification would have occurred had Newton simply explained Kepler's laws. The three laws would have been replaced by a single law, from which they are then rather complicatedly derived. However, Newtonian mechanics explained a tremendous number of phenomena, properties, and observations, thus offering the possibility of developing a mechanistic picture of the universe.

Another example is the development of the theory of electromagnetism which unites such phenomena as the propagation of light and the functioning of a dynamo. Maxwell's equations made possible a strict derivation of the laws governing the emission of radio waves, their reception by radio, the propagation of current through wire grids, and the operation of electric motors and transformers. No wonder that fairly complicated mathematical techniques were required. These are tools, merely tools. The whole physics of electromagnetism is contained, even if deeply buried, in Maxwell's equations which have a very simple form. They become especially simple if written not with obvious three-

dimensional coordinates but with more abstract although simpler four-dimensional coordinates in which time and spatial coordinates are replaced by the unifying concept of the "world point".

Or let us look at the diversity of matter in nature. The first simplification is the identification of about one hundred types of atoms of the chemical elements; these atoms are the basis of all the diverse forms of materials in the universe.

"If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the *atomic hypothesis* (or the *atomic fact*, or whatever your wish to call it) that *all things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another*. In that one sentence ... there is an *enormous* amount of information about the world, if just a little imagination and thinking are applied."*

Atoms, however, are not elementary particles. Nature has only three building blocks, not one hundred. Atoms consist of electrons, protons, and neutrons, which at present are considered elementary. This appears to be the most impressive simplification in the picture of the world ... .

But even the most perfect theory, one that describes an extremely wide scope of natural phenomena, is not perfect in the sense that it

* R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics*. Vol. 1, Ch. 1, pp. 1-2, Addison-Wesley, Reading, Mass., 1963.

does not include the limits of its own applicability. Only after a more general theory is worked out will it become clear what the bounds of the validity of the original theory are.

As a result, the possibilities of a theory are sometimes overestimated. Thus, the "triumphal advance" of Newtonian mechanics led to attempts to explain everything in terms of this mechanics only, and it was only the development of quantum mechanics that brought about an understanding of the futility of such attempts. It was also found that Newton's laws hold only while objects move at velocities small when compared with the speed of light. This restriction became clear after Albert Einstein had developed relativistic mechanics ... .

Theories are often said to follow, or replace, one another. This statement is both true and false. The development of a new theory does not invalidate the older theory* but makes the latter a particular (limiting) case of the new one (we have just mentioned that Newtonian mechanics is identical to Einsteinian mechanics when objects move at velocities small when compared with the speed of light).

What does it mean to comprehend the laws of nature?

Let us try to answer the following question: Are the laws of nature comprehended by man? The question is very difficult to answer since its formulation is not sufficiently rigorous. We shall begin at a remote point. We assume that

---

* *Note*: not everything referred to as a theory is indeed a theory.

a new phenomenon is discovered. We emphasize that it is a phenomenon. What is observed is not just something that cannot be explained immediately, but a previously unknown property of matter (of matter, not of the measuring instrument). It is found, for instance, that upon cooling almost to absolute zero some metals pass electric current without any resistance (superconductivity). As a rule, it is clear (or seems to be clear) which theory is to "govern" the new phenomenon.

Sometimes the phenomenon discovered has been theoretically predicted, that is, derived from fundamental equations by means of mathematical techniques. In the earlier history of science, a good fit of theory and experiment was regarded as something extraordinary. Such situations were described by such pompous phrases as "triumph of the power of thought", "discovery at the tip of a pen", and so on. Although in our time the confirmation of theory by experiment is much more frequent, the pleasure this confirmation brings is still acute, at least for the theorist (the expectation of such confirmation is possibly the most powerful stimulus for the theorist).

But sometimes it happens that the phenomenon was neither predicted nor can it be interpreted by the theorists after it has been discovered. This was precisely the case of superconductivity. What happens then? Is the reigning theory discarded? By no means! At the worst, one starts to doubt that the phenomenon really lies in the field of the theory to which it was assigned. But first of all one tries at least to give a correct *description* of the new phenomenon. "To describe" means to determine the mathematical relations

between the measured characteristics. Naturally, the theorist always tries to minimize the number of original equations and to maximize the number of corollaries derived from them. At this preliminary stage the theory usually includes special assumptions which are added in order to describe the new phenomenon. This produces a *phenomenological theory* (in the example of superconductivity, it was the theory developed by F. and H. London). The phenomenological theory stimulates new efforts, systematizes the data already accumulated, and points precisely to the relations in the theory which have to be explained by the general approach, that is, it indicates the type of formula which must be derived mathematically from the original fundamental equations; this way we can come to interpreting the whole set of experiments relevant to the new phenomenon. Once the formula has been derived, the phenomenon is finally understood, that is, a closed theory of the new phenomenon is completed. Thus, the theory of superconductivity was developed by three American scientists, J. Bardeen, J. Cooper, and J. Schrieffer, in 1957, that is, 45 years after the experimental discovery of the phenomenon. Therefore, understanding (or, in the rather high-flown style of the title of this subsection, comprehension) is equivalent in this case to the derivation of the required relation from the fundamentals of the general theory.

Let us return now to the problem of the degree to which the laws of nature are comprehended. In the sense of how we understand separate phenomena, the laws of nature are not yet comprehended.

We do not have, so far, unified (general) theory with respect to which all physical theories (mechanics, electrodynamics, theory of gravitation, etc.) are particular (limiting) cases.

Moreover, an element of description (phenomenology) is unavoidable. By definition, the fundamental (universal) theory can be formulated only on the basis of experiment, without reference to a more general theory. It is too early, however, to talk of the fundamental, or universal, theory. As yet it does not exist, but there is hope (or, rather, faith) that the universal theory will be constructed. Albert Einstein devoted the greater part of his creative life to a search for such a theory, which he called the *unified field theory*. But the lack of the unified field theory does not mean that physics is all chaos and disorder. Physics is the most logical of all natural sciences, in the sense that it operates with profound physical theories, which cover a tremendous number of properties and effects in nature, which are complete since they are internally noncontradictory and logically consistent, and which beautifully explain all the relevant experiments. Consequently, if the question: Are the laws of nature comprehended by man? is reformulated: Are the observed natural phenomena understood?, then the answer will be: Yes, they are for an enormous number of phenomena; this is definitely a triumph of human endeavour.

Where are the frontiers of knowledge?

The frontiers of the areas in which we have achieved comprehension have indeed pushed very far. These areas cover all observable phenomena in

the inorganic world: atoms and their nuclei, molecules in the form of macroscopic objects, that is, ensembles of 10^{23}-10^{24} particles, the phenomena within the solar system, the motion of star clusters, and the physical processes within stars. *Terra incognita* lies beyond this: in the properties of elementary particles and at distances of the order of 10^{10} light years away from us.

We are not claiming that each phenomenon in the above areas has been explained. We have already seen an example where it took nearly fifty years to reach an understanding of a discovered phenomenon. There are phenomena that still resist theorists' attacks. But we are certain that these will be finally explained within the framework of already known fundamental laws. Unless we go beyond the fantastically large area delineated above, there is no reason to reconsider the fundamental laws we have at present. Nevertheless, physicists acutely feel the lack of a unifying platform for physics as a whole; in the current language this means the failure to unify the theory of gravitation and quantum mechanics, to construct consistent relativistic quantum mechanics (which should include the theory of elementary particles). This explains why the branches of physics which enjoy maximum popularity (in the good sense of this word) are high-energy physics and the physics of elementary particles; this is physics which strives to go beyond the frontiers of the fields already mastered and thus to come closer to uncovering the very foundations of physics. This approach seems to be worth the "trouble" in spite of costs which average one billion (thousand million)

dollars and—this is the most important—in spite
of the high concentration of the human intel-
lect.

Micro- and macrotheories

Let us for the time leave the science of the future
and try to describe the state of physics at the
present moment. We shall step back and clarify
some of the points mentioned earlier.

The discovery of the atomic structure of mat-
ter, and later of the complex structure of atoms
and atomic nuclei, made possible the understand-
ing of macroscopic objects (more generally, of
objects with complex structures) on the basis
of the properties of microscopic objects. This
gave rise to the molecular kinetic theory of gases,
solid state theory, and other similar theories
in which the objects being described are macro-
scopic bodies, that is, enormous ensembles of
microscopic particles. Another, and equally im-
portant, feature is the following. When macro-
scopic objects are studied, we are not interested
in the motion (or states) of all particles compris-
ing the object, but want only to find some mean
characteristics: temperature, heat capacity, elec-
trical conductance, magnetic moment, and so
on. On the one hand, this self-imposed rejection
of extremely detailed (unnecessarily detailed)
description, and the introduction of statistical
(probabilistic) terminology are a corollary of the
impossibility in principle of any other approach;
on the other hand, this restriction follows from
the inadequacy of alternative approaches if a
correct description of macroscopic phenomena
is required. Say, in studying the behaviour of

a gas, we are interested in obtaining pressure as a function of temperature and do not want to record every impact of every molecule against the wall of the container.

This second approach in which every particle is "watched" does not invalidate the concept of "pressure" but clarifies it from the microscopic point of view, deciphers the concept, and demonstrates that the terms used in the macroscopic description can be reduced to more elementary concepts or events.

It might seem that this reduction of macroscopic theories to microscopic theories strips the former of their specific character. Or rather, that the macroscopic motion is stripped of its specific nature. But this is definitely wrong. For example, each microscopic motion is reversible while real macroscopic motions are irreversible*. Moreover, although macroscopic motions can be reduced to sequences of microscopic motions, the very nature of the irreversibility becomes clear from the procedure involved in this "reduction". The normal procedure for obtaining macroscopic equations is the averaging of microscopic equations. This averaging gives rise to the irreversibility of macroscopic motions, which is a qualitatively new feature when compared with microscopic motions.

These arguments about reducibility and specificity are very important because of the fre-

---

* Reversibility means that if all the particles involved in some motion suddenly reverse their velocities, their motions will be repeated in a reverse sequence. But a sled midway down a slope will not have the speed to climb the hill it was sliding down (a part of the energy was lost to overcome friction).

quent misunderstandings, in particular, those concerning the reducibility of biological processes to physical and chemical processes. The main argument against reducibility is the specific nature of biological processes. It must not be forgotten that, unless the debate is to be reduced to a useless terminological disagreement, specificity does not contradict reducibility. On the contrary, specificity can only be understood in the process of reduction.

Hierarchy in modern science

We have mentioned the relationship between macro- and microtheories. Let us emphasize one very general feature of modern physics, namely, its *hierarchical structure*.

A clear-cut division of science into subbranches does not contradict the fact that extremely interesting results often appear at the boundaries separating subbranches or even separate sciences.

The well-known microscopic structure of materials can be ignored when, for instance, technical problems of heat transfer or structural mechanics are being examined. An engineer designing diesel locomotives and dealing with various properties of metals (electrical, mechanical, and thermal) need not worry about electrons being charge carriers in metals. He is satisfied with using tabulated macroscopic characteristics of metals.

Obviously, special fields appear in physics not only when we move from micro- to macroobjects. It is quite typical for a particular case of a general theory to "break away" and form a spe-

cial field of science. This situation is illustrated clearly by, on the one hand, quasistatics (electro- and magnetostatics) and, on the other hand, radiophysics and optics, which have broken off the general theory of electromagnetism. This separation was justified by differences in the experimental and mathematical (i.e. theoretico-physical) methods of study in each of the sub-fields. The differences in techniques give rise not only to different terminology but also to different systems of images and concepts. Specialists in one field encounter difficulties in understanding their "neighbours". As a result, terms from the "neighbouring" field of physics are often introduced intentionally. For example, the properties of waveguides are "clarified" in terms of the theory of long transmission lines (in fact, this is an attempt to reduce a typically radiophysical problem to a quasistatic problem). Hierarchy is again "at work" here.

A physicist working in a narrow field of research may not refer to the general theory of which his field is a particular case (we are not going to discuss whether this is good or bad). This means that a corresponding limiting transition creates a logically closed science with its own equations and, consequently, solutions and conclusions. It would be virtually impossible to advance if we had to start at the universal starting line (*ab ovo*) for each new problem.

Abstraction and model

Among the characteristics of physics (and probably of any other science) we find one additional feature which is rarely emphasized but which

is absorbed naturally into the flesh and blood of every scientist. We mean the necessity of *abstraction*, that is, the necessity to find and single out among innumerable relationships only those which affect the relevant phenomenon to the maximum extent.

An abstraction materialized is a model. In reality, science deals only with models. And this statement is not aimed exclusively at theoretical physics. An experimenter studying collisions of particles in a gas tries to "get rid" of hundreds of disturbing factors which invariably affect real collisions. As a result, he manipulates a model of the gas, and the better the model is the less its gas resembles the contents of, say, a pressurized gas container installed to heat a building.

Fundamental theories (mechanics, electrodynamics) also resort to models. It seemed, at least until very recently, that nature's fundamental "building blocks" can be modelled by elementary particles (electrons, protons, neutrons, and neutrinos). In recent years, however, the number of elementary particles was found to be so great and their interaction so complicated (this can be seen in mutual transformations) that their properties hardly justify the term "elementary particles". One feels that what is required is another "reduction", that is, another penetration into the nature of elementary particles. This will not necessarily mean finding subparticles of which these elementary particles "consist" (in this sense they may prove really elementary) but of something different which at present cannot even be named.

What had to be sacrificed when quantum mechanics was worked out?

The appearance of quantum mechanics (it would be better to say the development of quantum mechanical philosophy) meant the breakdown of classical concepts. What specifically was demolished? In order to answer this question, it would be worthwhile to formulate briefly the main concepts of classical mechanics (obviously, a complete presentation is out of the question).

A material point which moves along a definite trajectory is the basic object of classical mechanics. The trajectory is determined by the initial conditions and by the force applied to this material point. It would seem logical that the discovery of the atomic nature of matter confirmed the fundamental concept of classical mechanics, with atoms, electrons, and protons—that is, the particles which must be treated as elementary in a specific kind of motion—being the material points (point objects). With this approach, the properties of macroscopic objects can be explained as the properties of ensembles of particles (point objects) that interact and move along definite trajectories. The mechanics of point objects can be constructed only if the forces of interaction between these objects are known. We shall discuss the forces later. We want to remark here that quantum mechanics does not revise the concept of force.

Summarizing, one can say that the classical mechanistic outlook is based on the concept of a material point, that is, a particle for which the coordinates and the velocity of motion can be determined at any moment of time. However,

the classical concepts are not restricted to this.

The discovery of radio waves and of the wave nature of light has shown that matter exists not only in the form of material particles (corpuscles) but also as waves. Habitually, mention of a wave is followed by indicating what moves in a wave-like manner. Indeed, if we speak about mechanical vibrations (for example, sound waves in gases), it is always clear that the objects involved in the oscillatory motion are particles (of the gas, liquid, or solid). But if we speak about electromagnetic oscillations, the concepts of the conventional model fail us. Nothing vibrates (if, by definition, "something" always means particles). In this sense the wave is a primary, irreducible concept. The electromagnetic wave (or, to be more precise, its most elementary form, a plane wave with a fixed frequency) is an elementary form of existence of a special sort of matter, namely the electromagnetic field. It is interesting to note that an elementary wave, that is, a wave from which all possible electromagnetic fields can be built, is a wave infinite in space and in time.

A wave and a particle appeared (not only at first glance but even, so it seemed, after a deep analysis) as incompatible, mutually exclusive concepts. Either we deal with a particle—something very small, which at any moment of time is at a definite point and is moving at a specific speed or with a wave—something continuous, filling up all space. To be correct, a wave is also characterized by a velocity, even two velocities. The first of them is the phase velocity, that is, the velocity of propagation of the phase of

the wave, $v_{ph} = \omega/k$, where ω is the circular frequency of the wave, and k is the wave number* related to the wavelength λ by the formula $k = 2\pi/\lambda$. The phase velocity characterizes the structure of the wave but not the velocity at which the energy of the wave is transferred. The second velocity is the velocity at which the energy of the wave travels, called the group velocity, since a wave train (group of waves) travels at this velocity. The group velocity is $v_{gr} = \partial\omega/\partial k$. If $v_{gr} = v_{ph}$, we say that dispersion is absent. For electromagnetic waves in a vacuum, $v_{gr} = v_{ph} = c \approx 3 \times 10^{10}$ cm/s.

The difference between a particle and a wave is especially pronounced when two waves or two particles are studied. A characteristic property of waves is the possibility of interference, which may result in diminished intensity produced by the superposition of two (or several) waves. For corpuscles this result appears absolutely unthinkable. Two particles are always more than one. In other words, the response to two particles arriving at one point must always be doubled (we assume for the sake of simplicity that the particles are absolutely identical).

We shall not discuss any more the obvious differences between waves and corpuscles, but shall make the point that sometimes this difference is not so obvious. When observing a beam of light, it is rather difficult to imagine that the "razor-sharp" beam is a superposition of infinitely extended light waves. One would rather agree that this is a stream of particles flying

---

* Wave number is the length of wave vector **k** directed along the direction of wave propagation.

along straight-line paths. We know that wave optics disposed of this difficulty with relative ease and found agreement between the rectilinear propagation of light and its wave nature by taking into account the interference of waves.

To conclude, we can say that the classical picture of the world is built of two elements: particles and waves.

A better expression can be found. The verb "build" seems too "heavy" when applied to waves. According to classical physics, the world consists of particles. These particles interact and move along certain paths. Some of the particles are charged. Moving with nonzero accelerations, these particles emit electromagnetic waves. This motion is perpetual and ubiquitous, so that the world composed of particles is filled with electromagnetic waves. Some of these waves are visible to us (visible light), but most of them travel through the ambient surroundings without affecting our senses. Some waves are easily picked up by radios or by other instruments that man has designed to broaden the range in which his senses are effective.

If we forget about the details, the above picture can be said to give a fairly complete representation of the surrounding world. Is it true to life?

Classical physics cannot explain the existence of atoms

The answer to the question posed in the last line of the preceding section is quite obvious. Were the classical picture adequate, there would be no need to speak of the revolution in natural sciences and no need to have the revolution itself.

But if something is said to be incorrect, like a theory not fitting an experiment, it is very important to know the character of the problem: Is the discrepancy between the experiment and the theory only slight, or is the theory very far from being in agreement with the experiment?

To what extent does the classical picture, outlined above, reflect the true nature of things? The answer will be ambiguous. As long as we are interested in macroscopic motions (or simply, motions of things, large objects), classical mechanics gives a very good description of experiments, and there is every reason to rely on its correctness. So long as the bounds of its applicability are not crossed (we shall soon indicate these bounds), nothing basic in it need be revised. But once we move into the world of atomic particles, this trustworthiness proves deceptive. Even the very existence of stable atoms and molecules cannot be understood if we stay within the framework of classical mechanics. Thus, we come not to a quantitative disagreement (of the type "the experiment gives 1.7 while the theory predicts 1.6") but to the principal impossibility of explaining the most fundamental facts. The importance of this is so great that, even at the risk of "forcing an open door", we will devote a few more sentences containing quantitative arguments.

Nowadays the size of the atom is well known. The radius of an "average" atom is approximately equal to $a \approx 3 \times 10^{-8}$ cm. It will presently become clear why 10^{-8} and not, for instance, 10^{-13} cm.

We can provide an answer to this question, although we have no idea why the proton is heav-

ier than the electron by a factor of 1840. True, we will not exactly *explain* why $a \approx 10^{-8}$ cm: we shall write a as a combination of some other quantities with well-known numerical values (known from experiment, of course). But that is all the explanation we need! In the case of the ratio 1840, however, we cannot express it via other quantities.

A positively charged nucleus "holds" a negatively charged electron by the force of Coulomb attraction*:

$$F_{\text{Coul}} = \frac{e^2}{a^2} \tag{1}$$

(e is the electron's and proton's charge; for the sake of simplicity, we consider the most primitive atom, that of hydrogen). For the atom to be stable, the force of attraction is insufficient (the electron would fall into the nucleus). This force must be balanced out by the corresponding force of repulsion. This role, as we know, is played by the centrifugal force:

$$F_{\text{c}} = \frac{mv^2}{a}. \tag{2}$$

Here m is the electron's mass, and v is its velocity. The equilibrium of forces (in the electron's centre-of-mass frame of reference) makes it possible to find the velocity of the electron in its

---

* Rutherford's famous experiments on the scattering of α-particles by nuclei constitute direct experimental proof of the validity of Coulomb's law at atomic distances. These experiments demonstrated that the force of interaction between the nucleus and the α-particle is inversely proportional to the square of the distance between them. Coulomb's law is violated only if this distance becomes of the order of, or less than, 10^{-13} cm.

circular orbit*:

$$v = \left(\frac{e^2}{ma}\right)^{1/2}.\tag{3}$$

With the numerical values of charge, the electron's mass, and the radius of the "average" atom, this yields $v \approx 10^8$ cm/s. On a terrestrial (and even cosmic) scale, this is an enormous velocity; but here it is essential to note that this velocity is three hundred times smaller than the speed of light. Hence, we may forget for a while relativistic effects. This conclusion is of paramount importance: the laws that operate within atomic distances are the laws of classical (that is, non-relativistic) mechanics. If the atom were as small as its nucleus (10^{-13} cm), the velocity of the electron would be close to the speed of light, in which case atomic mechanics would be relativistic from the very beginning.

We need two more formulas. The first gives the total energy of the electron in the field of the nucleus:

$$E = -\frac{e^2}{2a}.\tag{4}$$

This formula is easily obtained by summation of the kinetic and the potential energy of the electron. The minus sign means that we have chosen the electron's energy at an infinite distance from the nucleus as the reference level. Note that the energy diminishes as the distance a decreases.

The second formula is that for radiation intensity by a charge e moving with acceleration w

---

* To simplify the analysis, we will not try to find the exact solution for the motion of a charged particle in the Coulomb field, but will restrict the discussion to the motion in a circular orbit.

33

(in our case $w = v^2/a$):

$$I = \frac{2e^2w^2}{3c^3} = \frac{2e^2v^4}{3c^3a^2} \text{ erg/s.} \qquad (5)$$

If the electron emits I ergs each second, then it will lose all its energy over the time interval

$$\tau \approx \frac{a}{v} \left(\frac{c}{v} \right)^3 \qquad (6)$$

and will fall into the nucleus; τ is the classical electron lifetime in the orbit. Although this time is much longer than the period of revolution of the electron around the nucleus, a/v, because c/v is much greater than unity, it is still fantastically short:

$$\tau \approx 10^{-10} \text{ s.} \qquad (7)$$

This is the best demonstration of the total failure of classical physics. We have made use only of the absolutely obvious conclusions (the instability of any stationary structure of charged particles and the emission of radiation by a particle moving with acceleration) and concluded that atoms cannot exist! This is possibly the most impressive contradiction between classical physics and experiment. The new mechanics—viz. quantum (or wave) mechanics—was developed precisely to eliminate this contradiction; this mechanics is based on completely new concepts whose formulation demanded that our old illustrative concepts be scrapped.

In contrast to classical Newtonian mechanics and relativistic Einsteinian mechanics, quantum mechanics was developed not by a single man but by a pleiad of the outstanding physicists of our time: Max Planck, Niels Bohr, Louis de Broglie, Werner Heisenberg, Erwin Schrödinger, Wolfgang Pauli, Paul Dirac and Max Born.

Uncertainty principle

We shall not trace the history of the gradual construction of contradiction-free, logically consistent quantum mechanics by the concentrated effort of these great physicists. Let us "take the bull by the horns" and try to formulate the fundamental principles which make it possible to give a correct description of the motion of atomic particles (for instance, electrons).

When describing the microscopic world, we have to abandon the graphically descriptive concepts of classical mechanics. An atomic particle cannot be thought to move along a definite path. The concept of "path" ("trajectory") is essentially linked to the necessity of knowing at the same moment of time the exact values of both the coordinate of the particle and its velocity. But one of the fundamental principles of quantum mechanics states: *a particle cannot have simultaneously a definite coordinate and a definite velocity, or a definite coordinate and a definite momentum.** This statement can be expressed in the form of an inequality (the famous *Heisenberg uncertainty relation*):

$$\Delta x \, \Delta p_x \geqslant \frac{1}{2} \, \hbar. \tag{8}$$

How should we interpret this relation, which is often raised to the rank of a principle? The symbol Δx (or Δp_x) stands for the uncertainty in coordinate x (or in the projection of momentum p on the x-axis, p_x). The term "uncertainty in a physical quantity" means the following. Let an

---

* Instead of velocity **v**, quantum mechanics operates with momentum **p**.
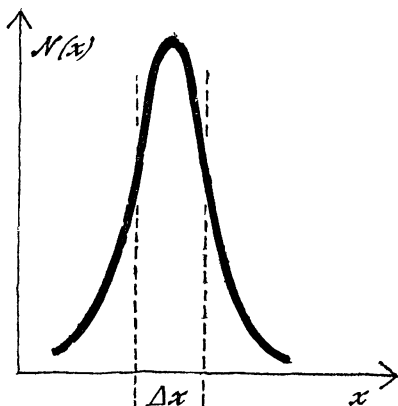
Fig. 1

atomic particle be in a fixed state (we should have
elaborated the term "state" as well, but we rely
on the reader's intuition; the words "a particle
is in a fixed state" can be said to mean "a particle
moves in a specified manner"). Very good, the
particle is in a fixed state. We measure its coor-
dinate and write down the result. Then we repeat
the measurement and record the result again,
performing this procedure many times. Now we
assign to each value of coordinate x the number
$N(x)$ which shows how many times this value
of x was recorded and plot $N(x)$ as a function of
x (Fig. 1). The curve clearly shows what is de-
noted by Δx. The uncertainty relation includes
a new parameter, a new letter $\hbar$; this is *Planck's
constant*, equal to approximately 1.05 $\times$
$\times 10^{-27}$ erg·s. In this book we write $\hbar$, in accord-
ance with long-standing tradition, although in
many books the preferred symbol is $h = 2\pi\hbar$.

Planck's constant is a new world constant. It arrived together with quantum mechanics when classical concepts had been discarded. Formally, a transition to classical mechanics is always possible by making $\hbar = 0$ (not that it is always easy, especially if we try to do it in the final formula). The Heisenberg uncertainty principle then indeed states that uncertainties in coordinates and momentum (Δx and Δp_x) can simultaneously equal zero, a particle thus being allowed to follow a trajectory.

Let us determine, without going deeper into the uncertainty principle, whether it imposes restrictions on the motion of macroscopic objects. Take a sphere with mass $m = 1$ g, whose trajectory is recorded within $\Delta x = 10^{-4}$ cm. What uncertainty in velocity is then given by the Heisenberg principle? The uncertainty relation yields:

$$\Delta v_x \geqslant \frac{\hbar}{2m\,\Delta x} \approx 10^{-23} \text{ cm/s}.$$

This example provides conclusive proof that there are no grounds to fear quantum uncertainty when considering the motion of a locomotive, or a car, or even their smallest parts. A molecule of air, striking a body with a mass of 1 g, causes a much higher change in velocity: $\Delta v \approx 3 \times 10^{-19}$ cm/s (we have assumed the molecule to be that of hydrogen with its mean energy at a temperature of the gas of 300 K).

Over the years, the uncertainty relation withstood numerous attacks launched from the most various sides. It is not a very well-known fact that the author of the theory of relativity Albert Einstein was one of the founding fathers of quantum theory and at the same time one of its staunch-

est opponents (paradoxically, the Nobel Prize was given to Einstein precisely for his contribution to quantum theory). With Einstein among the opponents of quantum theory, it is difficult to believe that the attacks stem from nothing more than the inertness of human nature, i.e. from its inertia in responding to anything new (although this factor definitely plays its role as well).

The uncertainty principle signifies that by its very nature the behaviour of an atomic particle must be described within the framework of the statistical (probabilistic) approach. It was the renunciation of the strict determinism inherent in classical mechanics that proved the most difficult to accept (even for Einstein!).

It is indeed very hard to imagine (a layman finds it difficult even today) that the behaviour of one electron which moves under the action of a definite force cannot be described completely without resorting to such notions as probability, uncertainty, or mean value. It seemed that these notions arise only in connection with a great number of particles. And this was wrong. The most rigorous analysis shows that the description of objects postulated in quantum mechanics is logically perfect and cannot in principle contradict the experimental data. To use a more precise expression, whatever phenomena, instruments, devices, or calculations are used, the momentum and coordinates of a particle cannot be determined with an accuracy higher than that imposed by the uncertainty relation.

It would seem at first glance that the uncertainty relation reflects the limitations of human capabilities: *we cannot determine* simultaneously

the coordinates and momentum of a particle with any better accuracy. But this is absolutely wrong! A *particle does not possess* the momentum and coordinates simultaneously. Any attempt to supply a particle with more accurate, the so-called internal, parameters invariably fails. Furthermore, it has been proved unambiguously that such attempts are incompatible with modern quantum mechanics, which is a science describing the world of atoms and molecules without any contradictions. This is surprising, especially if we take into account that attempts to assign dimensions to the electron, to find its radius, were equally futile. It should be emphasized that the coordinate uncertainty is an uncertainty in the *position* of a particle and is not connected with the size of the particle in any way.* In other words, an electron must be regarded as a point somewhere within an atom.

The uncertainty relation has an unexpected corollary which is especially important for solid state physics: the motion of a quantum particle (electron, ion, or atom) localized in a limited region of space cannot be stopped. Indeed, a stop ($p_x = 0$ and, hence, $\Delta p_x = 0$) would contradict the Heisenberg inequality since it demands complete delocalization of the particle ($\Delta x = \infty$). This unavoidable motion is often referred to as zero motion. We shall encounter it later.

---

* In some cases the electron's radius is mentioned. It is $\approx 10^{-13}$ cm, that is, smaller by a factor of 10^5 than the uncertainty in the electron's coordinate within an atom (see below).

Atomic size. Energy levels

Let us return to atomic particles. An electron in the atom moves along its circular orbit at a speed of $v \approx 10^8$ cm/s, that is, its velocity vector constantly changes the direction. On this basis we can assume that $\Delta v_x \approx v$. Indeed, if the electron's motion is projected onto the plane perpendicular to its trajectory, the velocity projection v_x oscillates from $-v$ to $+v$, with $v_x = 0$ at points of reversal. This is why we can say that the uncertainty in velocity Δv_x is of the order of the velocity itself. Hence, the uncertainty in coordinate of the electron, Δx, is not less than $\hbar/2mv$. The electron's mass $m \approx 10^{-27}$ g, so that $\Delta x \geqslant 10^{-8}$ cm and thus corresponds to the atomic radius. This means that a sphere of radius a encloses the volume in which the electron is definitely localized, but nothing more precise can be said about its position within the sphere. Taking into account that

$$a \approx \frac{\hbar}{mv}$$

and substituting instead of v the expression derived earlier, we obtain

$$a = a_0 = \frac{\hbar^2}{me^2}. \tag{9}$$

We have made good our promise: the atomic size a is expressed via other symbols—Planck's constant $\hbar$, the electron's mass m, and its charge e. This quantity, approximately equal to 0.5×10^{-8} cm, is called the *Bohr radius* and coincides with the radius of the hydrogen atom in the ground state. The hydrogen atom radius was first calculated by Niels Bohr, whose name is inseparable not only from numerous achieve-

ments of quantum physics but also from the development of the new philosophy in modern physics based on quantum mechanics.

Misunderstandings are quite possible when the fundamentals of quantum mechanics are presented in such a hasty manner. We want to warn the reader to beware of one of them.

In classical physics a nonradiating particle attracted to a "Coulomb" centre may orbit this centre along a number of trajectories whose shape and dimensions are determined by the integrals of motion (one of them is energy). The radius of a circular orbit depends only on the electron's energy; this dependence is given by eq. (4). After reading this section the reader might think that according to quantum mechanics an electron in the hydrogen atom can be only in the state whose classical model is the orbit with the Bohr radius $a = a_0$. This state (of course, the orbit is best forgotten) corresponds to energy $E = E_0 = -me^4/2\hbar^2 = -13.53$ eV (in atomic physics, energy is often measured in electron-volts, 1 eV $\approx 1.6 \times 10^{-12}$ erg). But this is not the only possible state of an electron in the hydrogen atom; this is the *ground state* (the minimum energy state).

In addition to the ground state, excited states are possible. But their energies are not arbitrary, as was the case in classical mechanics; only specific states are allowed. The energy of an electron in the hydrogen atom may be equal to

$$E_n = -\frac{me^4}{2\hbar^2}\frac{1}{n^2} \quad (n = 1, 2, 3, \ldots)$$

(Fig. 2). If the electron never emitted radiation, it would remain in any one of these excited states
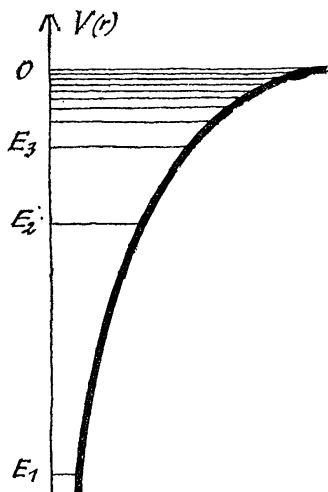
41

Fig. 2

perpetually. But here the impatient reader, given the chance, would interrupt the author: "If the electron *were not emitting*, it would exist eternally at a distance *a* from the nucleus in the classical atom as well! In both cases it would simply mean conservation of energy ... ." My reader, your remark is correct. But according to quantum mechanics only the states with energies E_n are allowed, and there is one state, namely the ground state, in which the electron *does not emit, ever.* This is argument number one. Second, the emission of a quantum electron is very different from that of a classical charged particle: the electron "jumps" from one level to another (from *n* to *m*), and the energy difference leaves the atom in the form of light. The law of energy

conservation definitely holds, but its form is an unusual one for classical physics:

$$E_n - E_m = \hbar\omega. \qquad (10)$$

Here ω is the frequency of the emitted light, and $\hbar$ is again the famous Planck constant from the uncertainty relation. This formula (also introduced by N. Bohr) explains the structure of atomic spectra which for many decades had seemed absolutely unfathomable. When the energy of an electron is comparatively high (n is large), the distance between neighbouring atoms is very small, and atomic "jumps" appear very much like a continuous approach of the electron to the nucleus. Calculations show that the frequency of the emitted light is equal to that of the classical rotation of the electron around the nucleus. The "classical electron" must emit light of exactly this frequency. This is one of the illustrations of the *correspondence principle*, that is, of quantum mechanics confirming a classical law. However, at the same time quantum mechanics indicates the physical range in which this law is valid.

The Pauli exclusion principle

We have solved in passing, while discussing the structure of the atom, one problem that puzzled Maxwell. We have found why all hydrogen atoms are identical: all of them are in an "identical" ground state.

One of the most important achievements of quantum mechanics was the explanation of the Mendeleev periodic law, in other words the explanation of similarities and differences among

chemical elements. A rather vague explanation can be given as: because different atoms have an unequal number of electrons*. For reasons of space, we cannot devote more than a few paragraphs to this topic. It would seem that the nucleus and its electrons generate a certain electric field within which the motion of each electron must be considered. This field contains "allowed" states, one of them being the lowest. This lowest-energy state "houses" all the electrons. Surely nothing will change if one electron is added (especially if Z is not too small). But this picture does not resemble reality at all: compare the properties of argon and potassium. One is a noble gas and the other is an alkaline metal, despite the fact that the number of electrons (the atomic number) differs by unity.

The structure of the atom cannot be understood without taking into account a very special selection rule** which was formulated by Wolfgang Pauli (1925). It is referred to as the *Pauli exclusion principle*. This selection rule forbids two (or more) electrons to be in identical states. To put it simply: one electron—one state (nature resembles a hotel with only single-room suites). We have to elaborate, though, what is meant by "state". Sometimes two states are united into

---

* The number of electrons within an atom (we denote it by Z) is equal to that of protons in the nucleus (see below) and is referred to as the atomic number.

** You have noticed that the word "law" belongs both to natural sciences and to jurisprudence. This is not a case of accidental homophony. Any law of nature, like any law in human society, is always a ban on something. For instance, a classical particle is not allowed to move along a trajectory if this trajectory does not follow from Newton's equations.
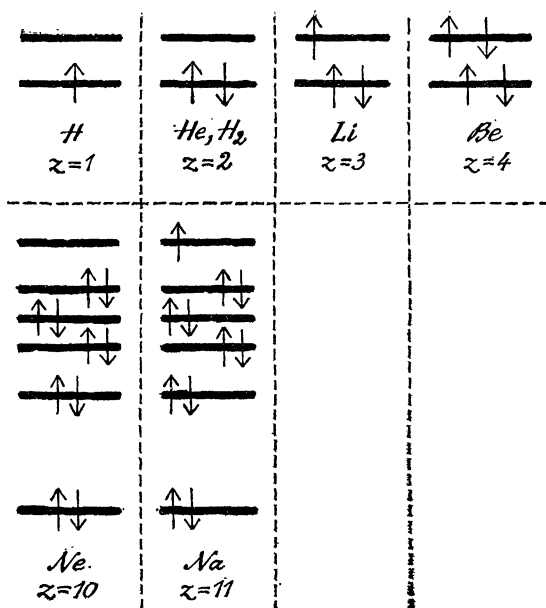
Fig. 3

one, with the ensuing modification of the formulation of the Pauli exclusion principle: two electrons—one state (a hotel with two-room suites). It should be added that not all particles in nature are subject to the Pauli exclusion principle.

This selection rule is so strict that it cannot be "beaten" even by the tendency, intrinsic to any physical system, to "roll down the slope", that is, to occupy the state with the lowest energy. The periodic law is thus explained by the distribution of electrons over the energy states of the

45

atoms according to the Pauli exclusion principle (Fig. 3). Thus, quantum mechanics becomes the foundation of chemistry.

Uncertainty relation and waves

It could be concluded from an analysis of the uncertainty relation that it is of a wave nature. Indeed, a formula of this type does hold in wave optics. We have already mentioned that a plane wave fills all space. It is well known, however, that the electromagnetic field can be concentrated within a finite (and if necessary, in a very small) region of space. This requires, however, that a large number of waves be used. Some of them enhance one another, while the interference of others results in their cancelling. This combination of waves is called a *wave packet* (or *wave train*). A wave train is characterized by an interval of wave numbers, Δk_x, required to concentrate the electromagnetic field in the region Δx. The set of wave vectors is the more varied the smaller the size Δx of the train. There exists a formula of purely wave origin:

$$\Delta x \, \Delta k_x \geqslant 1. \tag{11}$$

We recall that here Δx is the spatial length of the wave train ("uncertainty" in coordinates), and Δk_x is the "uncertainty" in the wave vector in the wave train. A comparison of the last inequality with the uncertainty relation shows that they are equivalent if we assume that

$$\mathbf{p} = \hbar \mathbf{k}. \tag{12}$$

This relation was first written by the French physicist Louis de Broglie in 1925 and is now known by his name. We can say that the uncer-

tainty principle means that particles have wave properties.

We have already mentioned that the uncertainty relation is a strict corollary of quantum mechanics. Quantum mechanics has also demonstrated, and just as rigorously, that waves have corpuscular properties and that particles have wave properties. The corpuscular properties of wave motion manifest themselves in the energy E of waves with frequency ω being equal to an integral number of quanta $\hbar\omega$ ($E = n\hbar\omega$, $n = 1, 2, 3, \ldots$), and in the momentum $\mathbf{p}$ of the wave being equal to $n\hbar\mathbf{k}$.

The wave-particle alternative (rather, the alternativeness of the concepts of "wave" and "particle") has been eroded at both ends: particles have gained wave properties, and waves have gained corpuscular properties. Both the wave properties of particles and the corpuscular properties of waves have been confirmed in numerous experiments, in which the electron, for example, behaves sometimes as a particle and sometimes as a wave. The same is true for electromagnetic waves: they behave as typical waves in some phenomena, and as pure particles in others. The picture of the world as traced in the preceding section totally contradicts our current concepts. The fundamental elements out of which we tried to construct the world have had to undergo revision.

The place of atomic physics in modern science

We have been discussing the relationships between the various sciences describing nature. And although the unified science, the "founda-

tion of all foundations", is still absent, it is clear from the point of view of physics that the primary sciences are those which deal with the most elementary objects—individual particles or waves. There are two such sciences: mechanics and electrodynamics (the field theory). Let us consider mechanics, that is, the science dealing with the elementary motions of particles. The separation from electrodynamics is certainly arbitrary since we remember that a charged particle moving with acceleration emits electromagnetic waves. At any rate, this is the picture if the motion of a charged particle is studied by using only the principles of classical physics.

Four mechanics are known at present:
(1) classical (Newtonian) mechanics;
(2) relativistic (Einsteinian) mechanics;
(3) nonrelativistic quantum mechanics;
(4) relativistic quantum mechanics.

The first three are logically closed sciences. The fourth is still being developed, not in the sense of obtaining new results but in the sense of completing the foundation, that is, the consistent logic of the science itself.

Each of the mechanics has its own range in which it is valid; the situation is conveniently illustrated by the diagram in Fig. 4. The variable laid off along the horizontal axis is the ratio of Planck's constant $\hbar$ to twice the particle's action S. The "action" is a quantity whose dimension is $g \cdot cm^2/s$, the same as that of Planck's constant. It is possible (and sometimes very convenient) to describe the motion of a particle in terms of its action. Quite strict rules exist to calculate the action. Suffice it to say here that S can be estimated by multiplying the charac-
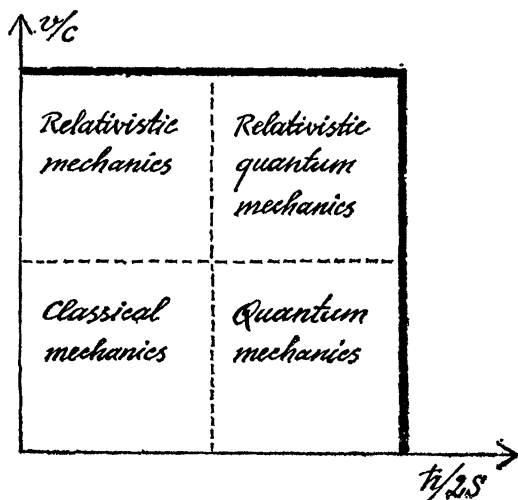
Fig. 4

teristic path covered by the particle in its speci-
fied motion by the characteristic momentum of
the particle. The variable laid off along the
vertical axis is the ratio of the particle's velocity
to the speed of light in a vacuum. The dashed
lines show the boundaries between the four me-
chanics.

Strictly speaking, relativistic or quantum laws
can be used to describe the motion of any object.
But is it necessary? The uncertainty in the ve-
locity of a sphere of 1 g mass has already been
calculated above. We have arrived at a figure
smaller than the change in the velocity of this
sphere caused by an impact of a single molecule.
Obviously, it would be unreasonable to intro-

duce quantum corrections in such cases. Hence, the position of the dashed lines in Fig. 4 shifts as the accuracy of calculation changes. There are at least two more reasons explaining the arbitrariness of the position of the dashed lines in the diagram.

There are phenomena which at first glance are attributed to the realm of classical physics. For instance, we would intuitively wish to classify magnetism, i.e. the capability of a magnet to attract iron filings or other iron objects, as belonging to the lower left-hand square. An analysis shows, however, that this would be erroneous, even if we assigned to this cell all the "super-structure"—macroscopic sciences which are generated by applying classical mechanics to an ensemble consisting of an enormous number of particles. Classical mechanics supplemented by statistics cannot explain magnetism in principle. This phenomenon is quantum by its nature. In a certain sense, this is even a relativistic phenomenon, although it operates with atomic particles; hence, it must be placed at the boundary between the two right-hand squares. It is not difficult to get rid of this ambiguity in the positioning of the dashed lines in the diagram. One must first interpret the phenomenon and then classify it under one of the squares of the diagram.

The second reason is much more complicated. We have already mentioned that three of the four squares represent *terra cognita*. These squares stand for logically closed sciences. But the fourth square comprises *terra incognita*, where surprises are in store for the scientist. Surprises not only in the sense of new discoveries but also

in the sense of new principles of constructing the scheme. We based our analysis on two fundamental constants—$\hbar$ and c. However, there are other parameters whose role is possibly not sufficiently clear at present. Maybe the diagram, more precisely the upper right-hand square, will have to be divided into smaller cells. For example, the electron's charge e, the speed of light c, and Planck's constant $\hbar$ can be arranged into a famous dimensionless combination:

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137}, \tag{13}$$

which is referred to very simply as "one over one hundred thirty-seven". It is also referred to, though less frequently, as the fine-structure constant. The inequality $\alpha \ll 1$ points to a weak coupling of particles to the electromagnetic field.* Other charges exist in addition to the electron's charge, for instance, the constant describing the interaction between nucleons (i.e. protons and neutrons in nuclei). The corresponding dimensionless combination is not small. In the current variant of the theory this leads to mathematical difficulties, complicates the techniques used by theoretical physics to describe the interaction between nucleons, but does

---

* At a certain stage, this enables us to treat particles and electromagnetic waves as the independent "building blocks" of which the world is composed. Note that the possibility of developing nonrelativistic quantum mechanics and of its application to electrons in atoms is in fact a corollary of the smallness of α. Otherwise it would not be possible to speak of allowed states in atoms. It is only owing to the smallness of α that the electron "survives" in the excited state for a comparatively long time and then "jumps" to a lower state. A special section will be devoted to 1/137 in this book.

not challenge the fundamentals of science. But maybe these fundamentals should be reconsidered? The revision may show that nuclear interactions must be classified as a separate region, with a third axis appearing on the diagram. So long as the building is not completed, it can be mentally organized in the most fantastic manner. Niels Bohr once remarked that the chances of success are the higher the more fantastic the approach.

One of the attempts to eliminate the difficulties* is based on introducing into the theory a fundamental length λ_0 or fundamental time interval $\tau_0 = \lambda_0/c$. Numerous such attempts are known. It is not easy to explain what is meant by "introducing λ_0 into the theory". In the simplest terms, it means that the structure of space-time must be revised. So far all the standard theories operate with continuous space and continuous time. The introduction of λ_0 signifies renouncing this continuity in favour of discrete space-time.

As follows from the available wealth of experimental data, the fundamental length, if it exists, must be of the order of (or less than) 10^{-13} cm. But this is the classical electron's radius! This value of the radius is found if the electrostatic energy of the electron with radius r_0, that is, $\sim e^2/r_0$, is made equal to the electron's rest energy mc^2. This gives

$$r_0 = \frac{e^2}{mc^2} \approx 2.5 \times 10^{-13} \text{ cm.}$$

---

* Difficulties must be understood as the inability of theory to give a consistent description of observed phenomena.

52

The equating of the electrostatic energy to the energy contained in the electron's mass is equivalent to assuming that the electron's mass as a whole is of electrostatic origin. So far we do not know whether this is true or not. But it cannot be excluded that many of the difficulties encountered by modern physics appear because particles are regarded as points. This means that their electrostatic energy (and mass) become infinite. Although formal methods overcoming this difficulty have been found, they have not yet been logically fitted into the fundamentals of the theory.

What are the components of the surrounding world?

We shall try to answer two questions:

(1) What are the components of the surrounding world?

(2) What are the forces? Later we shall clarify between what the forces act, but only after the first question is answered.

Strange as it may sound, the correct answer to the first question (rather, the beginning of the answer) must be another question: And who are you? The point is that the people working in the different fields of science expect different "depth of penetration" while answering the question: What are the components?

Biologists were satisfied for a long time with the answer: living matter consists of cells. Later one could hear: living matter consists of proteins, proteins consist of amino acids, and amino acids, of atoms and molecules. Molecular biologists are interested to know what atoms and molecules

are found in specific proteins. As a rule, a chemist is interested in the molecular composition of a substance, and only rarely in the atomic composition. Oxygen for a chemist is the molecule O_2. A physicist is not always interested in the components of atoms or molecules. For instance, if a physicist studies the condensed state of matter, he is more worried about the behaviour of molecules and atoms as such, and not about their structure. Could subjectivity in answering the question: What are the components? be avoided, or is the answer always biased?

In essence, this question means: How is it composed, or connected, and therefore what can it be decomposed into? Decomposition requires an expenditure of energy. What does a chemist do in this situation? As a rule, he dissolves a substance, sometimes heating it in the process. The energy transferred per particle is comparatively low. Therefore a chemist breaks only the weakest bonds. These are bonds between molecules. This is the reason why for a chemist matter consists of molecules. Sometimes this may not be so even at the "chemical level". If common salt (NaCl) is put into water, its dissolution is accompanied by electrolytic dissociation, that is, by decomposition of NaCl molecules into ions Na^+ and Cl^-. This means that common salt consists of positive Na^+ ions and negative Cl^- ions. The physical nature of the process of dissociation will be emphasized by remembering that dissociation is caused by decreased electrostatic (Coulomb) interaction between charges in the medium with dielectric constant ε. The force of attraction (and that of repulsion) is inversely proportional not only to the squared distance

but also to the dielectric constant ε. In water ε is comparatively high (ε ≈ 80), so that the force reduction is sufficient to achieve "breaking" of the molecule into individual ions.

A physicist can apply greater forces to a particle, for example, accelerate a projectile particle and make it hit the target atom that is being studied. The impact may remove an electron from the atom. We conclude then that an atom consists of electrons and a positively charged nucleus. The structure of the nucleus can be studied by using more energetic projectiles (such as accelerated protons). The most pertinent name for this procedure would be the "method of fragmentation". We have smashed the object to pieces and are sure that putting it together will reconstruct the whole and will give a correct idea about its components. But is the method of fragmentation so obviously correct?

An impact of the projectile particle does not necessarily ionize the atom (ionization means loss or gain of electrons), but may result in its excitation. The colliding projectile may transfer to the atom the energy $E_1 - E_0$, $E_2 - E_0$, etc., thus raising the atom to an excited state (see p. 41). In contrast to the *elastic* collisions in which the impinging particle changes only its direction of motion and the state of the target remains unaltered, the collision accompanied by the excitation of the target is called *inelastic*. The discovery of the step-by-step excitation of atoms played a most important role in the development of atomic physics.

An excited atom is thus produced, tending to return to the lowest-energy state. As a rule, this is achieved by the emission of the excessive

energy in the form of light. But does it mean that an excited atom consists of a light wave and the rest? Of course not! The light wave was generated at the moment when the excited atom was transformed into an atom in the ground state. Hence, an atom may decay into something of which it does not consist. But may it be that light waves are an exception in this respect? Indeed, how could an atom, whose size is only 10^{-8} cm, have enough "room" for a wave with the wavelength of, say, one centimetre?

Sometimes, when the projectile particle scores a direct hit on the nucleus (its radius is a hundred thousand times less than the atomic radius, so that such collisions are very rare), the nucleus is excited, that is, absorbs the energy of the projectile (inelastic scattering in the nucleus is said to have occurred). This is followed by emission from the nucleus. Different types of emission are possible. Sometimes the emitted particles are protons, sometimes neutrons, and sometimes simply light quanta (so-called γ-quanta, with the wavelength far beyond the upper limit of the visible light range). Sometimes these are electrons or their antiparticle counterparts, namely positrons, differing from electrons only in the sign of charge. This phenomenon (emission resulting from nuclear bombardment) was called *artificial (induced) radioactivity*. It was discovered by Frédéric Joliot-Curie.

The word "artificial" distinguishes this phenomenon from natural radioactivity, that is, spontaneous emission from the nuclei of heavy atoms (heavier than lead) of α-particles (nuclei of helium atoms), γ-quanta, and β-particles (electrons and positrons). An analysis of β-decay

has demonstrated that in addition to an electron or a positron, a nucleus emits one more particle: a neutrino (ν), a neutral particle with the zero rest mass. All these particles have been identified beyond any doubt: β-particles are indeed electrons or positrons. Does this mean that nuclei indeed contain electrons? This concept was prevalent for a comparatively long time, and the nucleus was regarded as consisting of protons and electrons. If the atomic mass of the nucleus is A, it would mean that the nucleus contains A protons and $A - Z$ electrons (Z denotes the atomic number*), that is, $2A - Z$ particles in each nucleus. With the development of a consistent theory, quantum mechanics, this hypothesis appeared less and less acceptable. But only the discovery of the neutron put everything in order. A neutron is a neutral particle whose mass is almost equal to that of the proton, and the nucleus contains Z protons and $A - Z$ neutrons, that is, A particles in all. Electrons and neutrinos, however, are produced at the moment of transition of the nucleus from one state to another. Therefore, β-decay must be interpreted as one of two reactions: $n \rightarrow p + e^- + \bar{\nu}$ or $p \rightarrow n + e^+ + \nu$, where $\bar{\nu}$ denotes the antineutrino. Note that only the first of these reactions can be observed in "pure" form, that is, outside the nucleus, since the proton is slightly lighter than the neutron.

Now we are able to give the final answer to the question: What are the components? Con-

___
* We have already mentioned that the atomic number coincides with the number of electrons in a neutral atom. Electrons are so much lighter than protons that their contribution to the nucleus' mass can be ignored.

densed matter consists of molecules and atoms, or of ions and electrons (metal). Molecules consist of atoms and ions; atoms, of electrons and nuclei; and nuclei, of protons and neutrons.

This simple scheme is not refuted by the latest discoveries in the physics of elementary particles. The main "building blocks" for matter are, as before, protons p, neutrons n, and electrons e. The remaining particles (whose number is uncomfortably great) are produced and then annihilated in nuclear reactions (as electrons and neutrinos are produced in β-decay). The picture of the world outlined above is not complete, but it has every right to exist since it gives a correct representation of the objective reality.

To express our profound respect for the "building blocks" of the universe, let us tabulate their individual characteristics.

	Charge	Mass	Magnetic moment	Spin	Statistics
p	$+1$	1838	2.89/1838	1/2	F
n	0	1838+2.4	1.9/1838	1/2	F
e	-1	1	-1	1/2	F

Particle charges in the table are given in units of the positron's charge, and masses are given in units of the electron's mass.

Spin and statistics of particles and antiparticles

The above table calls for some clarification. What are the strange units in which magnetic moment is given? What is meant by "spin" and

what is denoted by "F" in the column "statistics"?
Let us start with the spin.

It has been found that the particles in our list
(electron, proton, and neutron) cannot be regard-
ed as simply points or spheres. The closest one
could do would be to picture them as rotating
spheres, and rotating permanently at that. This
rotation can be neither decelerated nor accel-
erated. It is an intrinsic property of a particle.
Related to this rotation is an intrinsic angular
momentum equal to $\sqrt{1/2\ (1 + 1/2)}\ \hbar$. The quan-
tity $\sqrt{3/4}$ is given in the form $\sqrt{(1 + 3/4)}$ to
point to the formula $\sqrt{s\ (s + 1)}\ \hbar$. The quantity
s is said to be the spin of the particle. A particle
with spin s is a particle with angular momentum
$\sqrt{s\ (s + 1)}\ \hbar$. All the particles in the above table
have half-integral spin (the neutrino is also a
spin 1/2 particle).

Paul Dirac worked out the quantum mechanics
of spin 1/2 particles and proved that each such
particle with electric charge e possesses the mag-
netic moment

$$\mu = \frac{e\hbar}{2mc}. \tag{14}$$

This quantity, approximately equal to 10^{-20} erg/G,
is called the *Bohr magneton*. In our table the
magnetic moment of the electron is assumed to
be unity. To possess a magnetic moment means
to be a source of a *magnetic field*. To possess an
electric charge means to be a source of an *electric
field*. The derivation of the expression for the
Bohr magneton by Dirac meant the unification
of the electric and magnetic fields. The electro-
nic magnetic moment is a vector. But it is a very

strange vector: it can be oriented in space in only two ways, either along ... or against ... . But it remains unimportant along what and against what, unless the electron is in the magnetic field. The angular momentum is always oriented in $g = 2s + 1$ ways, so that for $s = 1/2$ it has two directions.

Each particle with nonzero charge generates a magnetic moment if its angular momentum $\mathscr{L} \neq 0$. The "classical" magnetic moment is $\mathscr{L}g_{cl}$, $g_{cl} = e/2mc$. Dirac's expression for the magnetic moment shows that g_q, that is, the proportionality factor between the spin momentum $(1/2)\hbar$ and the magnetic moment μ, is twice the classical value: $g_q = e/mc$. The quantity g_q is called the *gyromagnetic ratio*.

The electron shows a practically ideal agreement between Dirac's theory and experiment, but in the case of the proton the magnetic moment is greater by a factor of 2.89 than follows from the above formula. The magnetic moment of the neutron is given in the table in the unit called the Bohr magneton only for the sake of uniformity; indeed, according to the formula it should have been exactly zero, owing to the neutrality of the neutron.

Note: the good agreement of the magnetic moment of the electron with the Bohr magneton is another corollary of the smallness of the fine-structure constant α. The magnetic moment of the proton is anomalously high, and that of the neutron does not vanish because of the strong interaction between nucleons and mesons (see below).

An unexpected corollary of Dirac's theory is the existence of antiparticles. This conclusion

60

was staggering to many scientists. All particles must have "antiparticle counterparts", that is, there must exist (and do exist) particles identical in all respects but charge (they have charge of opposite sign) and magnetic moment (also reverse). But the most important point is the possibility of annihilation. Upon collision, a particle and an antiparticle disappear (annihilate), producing quanta of light ... . Conversely, a light quantum may produce two particles, for example, an electron and a positron ... .

There is an antiparticle to each of the particles mentioned so far:

electron—positron,
proton —antiproton,
neutron—antineutron,

and even neutrino—antineutrino.

Now that the properties of antiatoms are the subject of research, and the design of engines based on the transformation

matter + antimatter $\longrightarrow$ photons

is discussed seriously, the existence of antiparticles does not surprise anybody; but in 1932, when the "positive electron" was first observed on a photographic plate, this was interpreted as a triumph of theoretical physics. Indeed, Dirac had predicted its existence.

And finally, several remarks about the last column of the table. The letter "F" stands for the first letter of the name of one of the most famous physicists of the twentieth century, Enrico Fermi. Electrons, protons, and neutrons are often called *fermions*. The class of fermions in-

cludes all particles with half-integral spin. Particles with integral or zero spin (spin may be either integral or half-integral) are called *bosons* (from the name of the Indian physicist Jagadis Bose). This seemingly inconsequential difference between particles (the difference between integral and half-integral spin) results in that the behaviour of a system consisting of a great number of fermions differs radically from that of a system of bosons. Only fermions are governed by the Pauli exclusion principle, bosons being independent of it. Indeed, bosons tend to collectivize (to condense) in a single state. This property of bosons constitutes the *sine qua non* of quantum light sources, that is, lasers, and explains such phenomena as superfluidity and superconductivity.

Photon

One of the questions posed earlier has still not been answered: What are the forces? We cannot answer it yet. Before we do, we must make the acquaintance of one more member of the family of elementary particles—the photon.

In addition to the fundamental particles listed in the table, the world contains electromagnetic waves in considerable quantities. We have already said that owing to the uncertainty principle the difference between a particle and a wave is not as large as it might seem at first glance. We have even stated a rule for the particle-wave and wave-particle transformations. The corpuscular properties of an object are characterized by its momentum, and the wave properties, by its wave vector. They are related by the de Broglie

relation (12) that can be read both from left to right and from right to left.

Let us move on along the way to uniting the corpuscular and wave properties of particles. For this purpose, let us write corpuscular and wave formulas side by side, remembering that the motion of a particle is characterized not only by momentum **p** but also by energy E, and that a wave is described not only by its wave vector **k** but also by frequency ω (m is the particle's mass, and c is the speed of light).

Free particle	Wave in vacuum
$E^2 - c^2 p^2 = m^2 c^4$	$\omega^2 - c^2 k^2 = 0$

Looking at these relations attentively, we find that the differences disappear (at any rate, after multiplying the second of them by Planck's constant squared) if we make $E = \hbar\omega$ and assume that $m = 0$ for the electromagnetic wave.

A strict rule of quantum mechanics states that to each particle with energy E we can assign a wave with frequency ω. And, as always, the reverse is true: if there is a wave with frequency ω, one can consider a particle with energy $\hbar\omega$. In other words, a particle manifests its wave properties, and a wave, its corpuscular properties. Where do we notice these properties? For instance, a wave with frequency ω cannot have the energy below $\hbar\omega$. This is really resembling a particle: indeed, half a particle is impossible! But classical concepts state that the energy of

a wave is proportional to its squared amplitude and thus can be arbitrarily low.

Thus to an electromagnetic wave we can assign a particle. This particle is called the *photon*. From a layman's point of view it is a strange particle. Its mass is zero, while its velocity is always constant and equal to c. Its energy, however, is far from constant: it is the higher the higher its momentum (or the shorter the length of the electromagnetic wave). Indeed, the formulas given above yield a relation between the energy and the momentum of the photon:

$$E = cp. \tag{15}$$

It should not be thought that this relation reflects any great difference between a particle and a photon. Rather, in a sense it shows the proton's similarity to a particle, but not to an ordinary, *slow* particle: it resembles an ultrarelativistic particle, that is, a particle whose momentum p is much higher than mc ($p \gg mc$). The formula $E = \sqrt{c^2 p^2 + m^2 c^4}$ shows that for particles with very high momenta the energy is approximately equal to cp. Note that the usual formula

$$E = \frac{p^2}{2m} \tag{16}$$

also follows from the relativistic expression $E = \sqrt{c^2 p^2 + m^2 c^4}$ if we assume that $p \ll mc$. Then

$$E \approx mc^2 + \frac{p^2}{2m}.$$

It is not surprising that we arrive at the standard expression for the kinetic energy of a free particle, in which the term mc^2 is the rest energy

of the particle. In Newtonian mechanics this term is unimportant because energy is defined within an additive constant (energy can be measured with respect to any reference level).

We have supplemented the picture of the world by adding photons (quanta of the electromagnetic field, or particles with zero rest mass) to particles manifesting wave properties (electron, proton, and neutron).

It should also be mentioned that the number of neutrinos in the world is undoubtedly great.* Neutrinos are produced in nuclear reactions and have extremely long lifetimes, since they are very poorly absorbed by matter. The role of neutrinos in the structure of the world is still largely unknown.

The relation $\omega = ck$ between the frequency and wave number of an electromagnetic wave was neither invented nor postulated. It can be derived from Maxwell's equations for the strengths of the electric and magnetic fields. In electrodynamics Maxwell's equations play the same role as Newton's equations play in mechanics. Maxwell's equations for a vacuum describe the propagation of electromagnetic waves, but they have a solution only if the frequency and wave number of a wave are related by the formula $\omega = ck$.

Interaction between particles

Now we are ready to discuss the forces between the particles making up the world.

So far we have established that a structural

---

* The relation between the energy and momentum of the neutrino is the same as that of the photon since the neutrino is also a zero-mass particle.

unit of matter is a specific combination of particles, and its choice depends on how much energy we have at our disposal for realizing the dissociation, or decay, of this matter. However, the amount of energy required for dissociation is unambiguously related to the force of interaction between the components of the system. Consequently, the problems of structure and forces are closely interrelated. At each stage of breaking up matter we must simultaneously find the forces of interaction between the fragments. Obviously, the interaction between complex particles (such as atoms or molecules) must be explained in terms of the interaction between their components.

Let us begin the analysis of the forces of interaction with the well-known electrostatic Coulomb interaction between charged particles.

In order to get a feeling for the Coulomb forces, we consider an extremely impressive but hypothetical experiment. What would be the force of interaction between two spheres each having a volume of 1 cm^3 if the discrepancy between the charges would be one part in 10^9? In other words, let each sphere contain $10^{-9} \times 10^{23} = 10^{14}$ noncompensated electrons or protons. The resulting force of interaction between such spheres placed at a distance of 1 cm would be about 10 tons! It is not difficult to see that the world as a whole would be absolutely different were the charge equilibrium followed less strictly than it actually is. Note that the highly mobile electrons travelling between bodies act as the regulating factor which precludes the violation of the charge balance.

We want to analyze the interaction between

two charged particles in some more detail. Let their charges be e_1 and e_2. One of them, e_2, is now removed an infinite distance away. The first particle creates around itself an electric field whose potential φ is e_1/r. What does this mean? It means that if the second particle with charge e_2 is placed at a distance r from e_1, the force applied to e_2 will be equal to e_1e_2/r^2; it will be directed away from e_1 if the charges are of the same sign, or towards it if they are of opposite signs. No doubt, we could start with e_2 fixed, and "bring in" e_1 ... .

What was the purpose of introducing the field and its potential φ? The purpose was to underline that something real is present at the point where we placed the charge, and that this something acts on the charge. A more scientific expression of this idea is: the field is introduced to get rid of action-at-a-distance.

However, whether we introduce the field or not, the force applied to two charges is

$$F = \frac{e_1e_2}{r^2}.$$

This formula must surprise and shock the reader who is used to thinking that the speed of light c is the maximum velocity at which anything can travel. Indeed, what does it mean, a force which is a function of distance between particles? Will a particle feel the change in force at exactly the same moment as when the other particle is shifted? This contradicts the fundamental principle of Einsteinian mechanics, that is, the statement that the velocity of propagation of any signal is finite. We know perfectly well what caused this misunderstanding: an approxi-

mated relation (and Coulomb's law *is* approximate) was assumed to be absolutely correct. There are situations in which it does not yield a correct description of the interaction between charges. Which description is then in agreement with the theory of relativity? It is a description based on the concept of the electric, or rather electromagnetic, field. The force acting on a particle placed at a point of space is determined by the strength of the electric and magnetic fields at precisely this point. If one of the charges producing the field is shifted, the field in the vicinity of the charge will change; this will produce a wave of perturbation that reaches the second charge at a distance r after a time r/c. Through this mechanism the electromagnetic field ensures short-range action. The picture is actually more complicated owing to the effect of self-action, that is, the effect of the field on the particle generating this field. One has to be very careful in the relevant calculations.

In terms of classical concepts, the interaction of charged particles is represented by the following scheme:

particle $\longrightarrow$ electromagnetic field $\longrightarrow$ particle

and it is now clear that its quantum physics counterpart is

particle $\longrightarrow$ photon $\longrightarrow$ particle.

Describing this at greater length, we could say that a charged particle generates a photon that is absorbed by another particle. This is the mechanism which is responsible for the

68

force acting on these particles. The reader will have to accept without proof that Coulomb's law (the force is inversely proportional to the square of the distance between particles) is the corollary of the fact that the photon has a zero mass.

The natural question is: What would the law be if the particles mediating the interactions possessed the finite mass? It can be shown that in such a case the forces of interaction would fall away more strongly than the Coulomb force. Even the form of force F as a function of distance r can be given for large r:

$$F \sim \exp\left(-\frac{r}{r_0}\right).$$

The quantity r_0 is called the radius of action of force F. It is related in a simple manner to the mass m of the mediating particle: $r_0 = \hbar/mc$.

We know of forces which rapidly diminish with distance. These are the nuclear forces which hold nucleons (protons and neutrons) together within a nucleus. These forces are very great (about one hundred times greater than Coulomb forces) but they are effective only at very short distances: their radius of action is approximately equal to 10^{-13} cm.

The Japanese physicist Hideki Yukawa conjectured that forces between nucleons are generated by particles which he called *mesons*. Why intermediate? We readily find from the last formula that for $r_0 \approx 10^{-13}$ cm the mass of the mediating particle must be approximately equal to three hundred electron masses, somewhere between the masses of the electron and proton. Hence, the term "meson".

The most exciting thing about this hypothesis was that the mesons responsible for nuclear interactions were indeed found after Yukawa had predicted them. They were called π-*mesons*, or *pions*. Three types of pions are known: π^-, π^+, and π^0. The superscript indicates the sign of the electric charge of the meson. Thus, the particle responsible for the coupling of nucleons is electrically charged. Charged mesons (π^- and π^+) can emit and absorb photons. The picture is becoming very complicated even for the interaction of elementary particles.

Pions are zero-spin particles, and therefore bosons. This information is supplied to guard the reader against any hasty conclusions. The table of elementary particles on p. 58 could give one the impression that all elementary particles are fermions. The photon is also a boson; but not all particles mediating interactions are invariably bosons. By no means; the weak interaction between nucleons is mediated by electrons and neutrinos (i.e. by fermions).

Summarizing, forces acting at distances of the order of 10^{-13} cm and smaller are nuclear forces coupling nucleons (protons and neutrons). From the point of view of nuclear forces, the neutron possesses a "charge" while the electron is "neutral", that is, it does not interact with the nucleon via nuclear forces even if it is separated from the nucleon by 10^{-13} cm.

Another important property of nuclear forces is their dependence on the spins of nucleons, that is, on the mutual orientation of spins. Indeed, spins resemble "magnetic needles" which can be oriented with respect to one another in different ways.

The next unit in the structural hierarchy of matter, after the nucleus, is the atom. Electrons in the atom interact with one another and with the nucleus. The main force within the atom is the Coulomb force. Our earlier estimate of the velocity of the electron within the atom demonstrated that it is small (by a factor exceeding one hundred) when compared with the speed of light. This fact makes it possible to ignore retardation effects and use, as a first approximation, the long-range Coulomb law.

Electrons, protons, and neutrons possess magnetic moments which also interact (magnetic poles of like polarity are repelled, while those of opposite polarities are attracted to each other). Do we have to take these forces into account? Or, perhaps, these forces do not operate at the atomic level? But, oh yes, they do. Not only do they operate within the atom, but the law describing the interaction of two microscopic magnets is the same as that for ordinary magnetic needles (if the needles are separated by a distance far greater than the length of each needle). The force between two small magnets, F_m, is inversely proportional to the fourth power of the distance between the moments and directly proportional to the product of the magnetic moments:

$$F_m \approx \frac{\mu^2}{r^4},$$

where μ is the Bohr magneton. (For the sake of simplicity, we take into account only the interaction between two electrons, the magnetic moment of the nucleus being much smaller; the approximate equality is used because F_m^{12} is a function not only of the distance but also

71

of the mutual orientation of the magnetic "needles".) The distance between electrons in an atom coincides, by order of magnitude, with the atomic size, a, so that $F_{\mathrm{m}} \approx \mu^2/a^4$, and the Coulomb force of interaction between particles is $F_{\mathrm{Coul}} = e^2/a^2$ (see eq. (1)).

By taking into account the expression for the Bohr magneton μ (see eq. (14)) and for the atomic size a (see eq. (9)), we find that magnetic forces are much weaker than electrostatic forces:

$$\frac{F_{\mathrm{m}}}{F_{\mathrm{Coul}}} \approx \left(\frac{e^2}{\hbar c} \right)^2 = \left(\frac{1}{137} \right)^2 \ll 1. \qquad (17)$$

This inequality is a corollary of ignoring magnetic interactions in crude approximations of the atomic structure.* The corresponding forces must not be forgotten since they play a significant role in a number of subtle but perfectly observable phenomena.

Let us move one more step up the ladder of increasing complexity of systems formed by microscopic particles. The forces between atoms are predominantly of electrostatic origin. Although an atom as a whole is electrically neutral, the charges in it are not localized at a single point. The following very rough model can be constructed: when atoms approach, charges of opposite signs are attracted, and those of like sign are repelled. The force applied to each atom is the difference between the forces of attraction and repulsion. The influence of one atom on the motion of electrons in another one is such that the resulting force is always

---

* Note that this is another manifestation of the smallness of the fine-structure constant.

72

attractive. This occurs until the two atoms almost touch each other. At such distances repulsion "steps in". If at a particular distance the force of attraction is equal to that of repulsion, a molecule is formed.

But what causes this force of repulsion? Once we established that the electron is about a hundred thousand times smaller than the atom, it seems logical to regard the atom as consisting predominantly of "vacuum", with the nucleus and electrons taking up a surprisingly small space within it. What, then, makes the atom behave as if it were a rigid sphere? Even serious monographs use the concept of the atomic radius. A crystal can (and must) be pictured as an arrangement of regularly placed balls (like billiard balls put in a triangle on the table in preparation for a game of pool).

When atoms move closer to one another, the "Lebensraum" for the electrons diminishes, and the uncertainty in the electron's coordinates also diminishes. This increases the uncertainty in the electron's momentum (recall the Heisenberg relation), and therefore their momentum as well. This means that the kinetic energy of motion of electrons rises, increasing the total energy of the system. But if the total energy of the approaching particles increases, it means that they undergo repulsion: it is advantageous to be at a distance, since this diminishes the energy ... . As you see, the force of repulsion has a specifically quantum origin.

There is no need for us to climb the ladder still further by analyzing the interaction between molecules. In principle, it is the same as in the case of atoms.

73

In this chapter we have mentioned several times the fine-structure constant, 1/137, which is, essentially, a *measure* of the electron's charge (provided the quantum of action, i.e. Planck's constant, $\hbar$, and the speed of light, c, are considered to be more "fundamental" than the charge: $e^2/\hbar c = 1/137$).

Let us recall two formulas: one for the size of the hydrogen atom, $a_0 = \hbar^2/me^2$, and another for the characteristic energy of the electron in the atom, $|E_0| = e^4 m/2\hbar^2$. They can be rewritten in the form:

$$a_0 = 137^2 r_0, \qquad |E_0| = \frac{mc^2}{2\,(137)^2},$$

$$r_0 = \frac{e^2}{mc^2} = 2.5 \times 10^{-13} \text{ cm}. \qquad (18)$$

The first formula explains the "emptiness" of the atom, while the second states the nonrelativistic character of the mechanics of the atom. It is interesting to note that sometimes the smallness of α manifests itself in an unexpected manner. For example, diamagnetics are never "strong": the diamagnetic susceptibility χ is always low. It can be shown that the value of χ is proportional to $(1/137)^2$. Numerous such examples are known. This section is intended to attract the reader's attention to a fact of principal importance. The existence of a dimensionless combination uniting quantities of very dissimilar nature is an extremely important feature which shapes the world into what it is.

74

I remember a chapter in Ya. Perelman s *Physics for Entertainment* giving a description of a 100%-relativistic world; in other words, a picture of our world with the speed of light greatly reduced and the rest unaltered (including our perception of the world). But what would a world be like in which $e^2/\hbar c = 1$? Or in which the ratio of the mass of the proton to that of the electron differed from 1840?

Particles or waves?

The particle-wave dualism, or duality, of atomic particles is a property so important for understanding the surrounding world that we must discuss in detail the experimental proofs of this striking characteristic of microscopic objects.

The proof of the dual nature of atomic particles and light is essentially simple and a basic knowledge of the fundamental concepts of physics is sufficient to understand it. The reader must also have some feel of the character of the behaviour of a wave and of a particle under specific conditions.

The duality of light had been proved earlier than the wave properties of the electron were discovered. The fact that the first quantum concepts were developed for light (they concerned its granular, discrete structure) is explained not only by the pressure of the experimental facts (observation is often separated from explanation by many years) but also by several centuries of struggle between the supporters of the corpuscular theory of light and the advocates of the

wave theory. The wave theory appeared as the winner of that struggle.

No one could doubt this outcome at the end of the nineteenth century. The wave theory was successful not only in explaining the propagation of light along straight lines and the laws of reflection and refraction (the corpuscular theory had its own interpretations of these phenomena, although with several arbitrary assumptions) but also in predicting the results of specially conducted experiments which were to give an unambiguous answer to the particle-wave argument. The answer was absolutely unequivocal—waves. We refer to the experiments on interference and diffraction. Furthermore, the wavelength of light waves was measured, and the correspondence between colour and wavelength was established. And if any doubts still remained they were finally erased by the development of the theory of electromagnetism. The electromagnetic nature of light waves was established unambiguously, the speed of the propagation of light was explained with spectacular success. The unification by Maxwell of electricity, magnetism, and optics into an elegant, logically consistent theory was a dramatic achievement of the human mind. It is an achievement that mankind remembers even when a more advanced theory has set bounds on the applicability of the older theory that before seemed all-embracing.

"From a long view of the history of mankind—seen from, say, ten thousand years from now—there can be little doubt that the most significant event of the nineteenth century will be judged as Maxwell's discovery of the laws of electrodynamics. The American Civil War will

pale into provincial insignificance in comparison with this important scientific event of the same decade."*

Hence, the wave nature of light has been established unambiguously. But there are several points we would like to dwell on here.

Blackbody radiation is the radiation of a body that absorbs all incident light but reflects nothing, or almost nothing. It is very easy to design such a radiation source. One has to arrange a cavity opening to the surrounding space through a small hole. This opening is the blackbody. A ray of light entering it will penetrate into the cavity, will follow a tortuous path, and will be absorbed before it is able to leave through the hole. But if the walls of the cavity are heated sufficiently, the cavity will certainly emit light; everybody thought that the emitted radiation should be light waves. Blackbody radiation was studied theoretically by Max Planck at the beginning of the twentieth century.

Physicists were very interested in the laws governing blackbody radiation (some of the laws had been discovered before Planck began his studies). It was in fact clear from the most general viewpoint that the laws of blackbody radiation are related only to the nature of the radiation itself and have nothing to do with the processes on the walls. An analysis of blackbody radiation enables us to study light (electromagnetic waves?!) in its pure form. It is important to underline the utmost simplicity of the object of study: light alone nothing else. The theory of

---

* R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics.* Vol. 2, pp. 1-11. *Electricity and Magnetism.* Addison-Wesley, Reading, Mass., 1964.

blackbody radiation was already available in Planck's time. It was based on concepts which were completely natural for the wave approach: electromagnetic waves are confined within the cavity. The cavity walls have a definite temperature. The waves must be in equilibrium with the walls and therefore have the same temperature T (as two bodies brought into contact).

A logical question would be: Are we justified in speaking about the temperature of waves? Yes, we are: the extension of the concept of temperature to include electromagnetic waves is perfectly natural (unfortunately, we cannot discuss this generalization here). The calculation of the density of electromagnetic radiation, E_ω, for a narrow range of frequencies yields

$$E_\omega = \frac{V\omega^2 k_{\mathrm{B}} T}{\pi^2 c^3},$$

where V is the cavity volume, ω is the light frequency, c is the speed of light, T is the temperature in kelvins, and k_{B} is the Boltzmann constant approximately equal to 1.4×10^{-16} erg/K; k_{B} "transforms" degrees into ergs, that is, gives temperatures in energy units. This formula is called the *Rayleigh-Jeans law*. It fits well with experimental data while we are dealing with low-frequency radiation, but fails completely if we are interested in the description of the whole radiation spectrum (of all frequencies).

One immediately notices that an attempt to determine the total energy emitted at all frequencies by using the Rayleigh-Jeans formula yields an absurd result. The total energy radiated is infinite. This situation was called the *ultraviolet catastrophe*. The experiments showed, in fact,

that at high frequencies radiation is not described by the Rayleigh-Jeans formula but by a strange expression (found when physicists tried to tie up the loose ends) that was in obvious contradiction with the classical picture: the energy density decreased exponentially as frequency increased. Max Planck thoroughly analyzed the very foundations of the derivation of the Rayleigh-Jeans formula and came to the conclusion that in the framework of standard classical physics nothing can be changed or corrected. Besides, no agreement is possible between this formula and the experimental data. Planck suggested a seemingly outrageous hypothesis: the energy of a wave with frequency ω cannot be arbitrary but assumes only discrete values $\hbar\omega$, $2\hbar\omega$, $3\hbar\omega$, etc. Later the constant $\hbar$ was called Planck's constant. Its value was given earlier in the text. This strange conjecture immediately led to a meaningful pattern. The formula derived by Planck (the *Planck radiation formula*),

$$E_\omega = \frac{V}{\pi^2 c^3} \frac{\hbar\omega^3}{\exp{(\hbar\omega/k_B T - 1)}}, \qquad (19)$$

fits perfectly with experimental results. When $\hbar\omega$ (the amount of energy equal to $\hbar\omega$ was called a *quantum of energy*, or simply *quantum*) is much smaller than temperature (in ergs), then the Planck radiation formula becomes identical to the Rayleigh-Jeans formula. Is a quantum of energy, $\hbar\omega$, a physically meaningful entity? This was rather difficult to fathom since within the framework of classical physics it was absolutely impossible to explain how an electromagnetic wave can have only prescribed discrete values of energy. And that was the only way to inter-

pret blackbody radiation. The physicists were utterly confused. It is clear to us nowadays that Planck's "crazy" hypothesis marked the beginning of the revolution in physics, the first "quanta offensive".

We shall note, before analyzing other phenomena, that the Planck radiation formula can be rewritten in the form

$$E_\omega = \frac{V}{\pi^2 c^3}\, \omega^2\, \frac{\hbar\omega}{\exp\,(\hbar\omega/k_\mathrm{B}T - 1)}\,. \qquad (20)$$

We notice that the factor $k_\mathrm{B}T$ is replaced by the expression $\hbar\omega/[\exp\,(\hbar\omega/k_\mathrm{B}T - 1)]$. This notation will be useful to us later.

The *photoelectric effect* is a comparatively simple phenomenon. The illumination of the surface of a metal can cause the ejection of electrons out of the metal. Careful studies of these electrons have demonstrated that their behaviour is rather strange. For some reason the energy of the electrons was found to depend on the frequency of the incident light, while the light intensity only increased the number of emitted electrons. It is not difficult to see that if light is an electromagnetic wave, then the energy of the emitted electrons is determined by none other than the intensity of the wave, that is, by its amplitude, and that the frequency is of no consequence.

The true nature of the photoelectric effect was realized by Albert Einstein. Light is an ensemble of quanta, or photons. The interaction between electrons in the metal and light develops as the absorption of light quanta (portions of electromagnetic energy) by electrons which are therefore ejected from the metal. The higher the frequency the higher the quantum energy. Therefore the

higher the frequency of light the higher the energy of an emitted electron. An increase in intensity means that the number of photons incident on the metal is greater, the number of absorbed quanta increases, and therefore the number of emitted electrons rises. Once again (as in the case of blackbody radiation) the hypothesis of the discrete, quantum structure of light leads to an excellent explanation of the experimental facts.

For our purposes it is better to rephrase this evaluation (into a sort of inverted theorem): blackbody radiation and experiments on the photoelectric effect demonstrate that light has a discrete, quantum structure, that is, consists of photons.

The Compton effect. The experiment that shows this structure most clearly involves the scattering of hard X-ray emission. There is no doubt that X-rays are waves. They are used to measure interatomic spacing in the crystal lattice (by comparing it to their *wavelength*). What does classical theory predict? First of all, scattered radiation is expected to have the same frequency (the same wavelength) as incident radiation. Indeed, the electromagnetic wave forces electrons to oscillate at the frequency of the wave and hence to emit waves of the same frequency. And what has quantum theory to say about the matter? According to quantum concepts, light is an ensemble of quanta, or photons, that is, particles with energy $\hbar\omega$ and momentum $\mathbf{p} = \hbar\mathbf{k}$ ($k = 2\pi/\lambda$, where λ is the light wavelength). A photon colliding with an electron transfers to it part of the energy and momentum. The loss in energy will be seen as a change in frequency or

wavelength ($\omega = 2\pi c/\lambda$ or $\varepsilon = cp$). In calculating this change of wavelength, $\Delta\lambda$, we need nothing but the laws of energy and momentum conservation. The final result is strikingly simple:

$$\Delta\lambda = \frac{2\pi\hbar}{mc}\,(1 - \cos\varphi), \qquad (21)$$

where φ is the scattering angle, that is, the angle between the directions of propagation of the incident and scattered photons; the quantity $2\pi\hbar/mc$ has the dimension of length and equals ≈ 0.024 Å; it is called the *Compton wavelength*, after Arthur Compton who measured the shift in the X-ray wavelength in 1922-1923 and found that it is in complete agreement with quantum theory.

When proving the discrete, quantum nature of light, we do not reconsider the results of the older classical experiments, such as the experiments on diffraction or interference of light. They need no reconsidering: they are correct. Equally correct are the other experiments that we described above. The remarkable fact is that there are some experiments that cannot be understood within the framework of wave theory, while there are other experiments for which corpuscular theory fails. Sometimes light behaves as a particle and sometimes as a wave. This does not mean that light is "free to choose" and behaves as it pleases. Modern quantum theory enables one to predict precisely which experiment will reveal the corpuscular properties of light and which will bring out its wave properties.

It is difficult, once you get used to the wave interpretation of optical phenomena, to get ac-

customed to the quantum interpretation. But by no means can we ignore it. It has been found that the granular structure of light can even be detected visually. The experiments of S. I. Vavilov demonstrated that a human eye well adapted to darkness is capable of resolving several separate quanta.

It is hardly necessary to persuade the reader that the electron is a particle. There is no basis to doubt that the corpuscular nature of electricity is possible after the discovery of the discreteness of the electric charge, the measurement of the charge and mass of the electron, the demonstration of traces of electrons in the Wilson cloud chamber, the measurement of the electron range in photoemulsions, the construction of gigantic machines for the acceleration of electrons, and finally, after electrons were used as projectiles for the bombardment of nuclei. We became accustomed to the electron and almost neglect the fact that it serves us in our daily life. Watching the adventures of the hero on a TV screen, we tend to forget that the animated pictures are traced for us by a beam of electrons.

Nevertheless, the wave properties of the electron are revealed just as clearly. Louis de Broglie was the first to arrive at a brilliant hypothesis that particles should have wave properties. We have mentioned this earlier. De Broglie conducted no experiments whose outcome would indicate this conclusion. He pondered over a surprising fact that equations of mechanics can be written in a form identical to those of optics. The similarity becomes complete if the energy of a particle is connected with a frequency, and momentum, with the wave vector of a wave proc-

ess. It took some time to realize that the proposed strange equalities

$$\varepsilon = \hbar\omega, \quad \mathbf{p} = \hbar\mathbf{k} \qquad (22)$$

spell a revolution in physics.

However Einstein wrote in his foreword to de Broglie's book *Physics and Microphysics* that de Broglie was the first to realize the closest physical and formal interrelation between quantum states of matter and resonance phenomena at the time when the wave nature of matter had not been yet discovered experimentally.

Fortunately, de Broglie's paper attracted the attention of the American physicists C. Davisson and L. Germer, who were studying the scattering of electrons by the surface of metals. As is often the case, a momentous discovery started with an annoying accident*: the vessel in which a nickel plate was kept (the reflection of electrons from the nickel surface was being studied) had broken. The plate had oxidized and had to be annealed for a considerable time. The results obtained after annealing differed sharply from those recorded before this procedure: the dependence of the number of scattered electrons on the scattering angle was found to be strongly anisotropic. The experimenters spent several years trying to decipher the puzzle. The explanation was obtained only when de Broglie's idea was used.

Here is what happened: annealing of the plate led to recrystallization, so that large crystallites

---

* Many authors like to emphasize the accidental character of discoveries. What is not accidental, though, is that these lucky accidents happen to those who are capable of understanding their significance.

were formed. These crystallites were responsible for the anisotropy of scattering. This anisotropy could be deciphered only by assuming that electrons are reflected by the crystal exactly as waves would. It was then necessary to accept that the wavelength of the electron satisfies the de Broglie equation ($\lambda = 2\pi\hbar/p$). Electrons are diffracted as a result of the reflection of electron waves from the crystal lattice planes, just as X-rays are.

The diffraction of electrons (a queer combination of words if the electron is assumed to be a particle) has become one of the basic methods of crystal analysis, in addition to the X-ray structural analysis and, no longer surprising, the diffraction of neutrons.

The discovery of the diffraction of electrons did not invalidate the corpuscular properties of the electron. These properties cannot be refuted—they exist. The wave properties, however, exist as well. The electron is a particle and a wave at the same time. This is the nature of microscopic particles.

Sometimes the dual nature of microscopic particles manifests itself in a single experiment (strictly speaking, in two simultaneous experiments). Take the scattering of a neutron by a crystal. One of two events will take place: the neutron will either conserve its incident energy or it will be somewhat slowed down.* The energy-conserving scattering (elastic scattering) is a clear sign of diffraction and as has already been mentioned is a reliable method for the structural

* At nonzero temperatures of the crystal, neutrons may even be accelerated, by taking energy from the crystal.

85

investigation of crystals (especially magnetic crystals, if the elastic scattering of neutrons is used), while in any description of inelastic scattering the neutron must be considered as a particle.

The Schrödinger equation. ψ-function

When we mention photons, we clearly understand that what is meant is a concept of portions of the (electromagnetic) energy of light. The question of the components of the photon simply does not arise. Neither is there any basis (at least at the present moment) to ask this question about the electron. This would contradict the concept of the elementary character of the electron.

But what is the wave process which accompanies the motion of microscopic particles according to the de Broglie relation. We have stated and proved on many occasions that particles possess wave properties. Very well, the electron is a wave. But, a wave is something extended both in space and time. In what way does a point in space in which a wave is present at a given moment of time differ from a point where there is no wave? In the last phrase we should really replace the word "wave" by the word "something": In what way does a point in space in which "something" is present at a given moment of time differ from a point where there is nothing?

It is customary to denote the quantity that oscillates in this wave by the letter ψ. If ψ at one point is zero, then there is nothing at this point. If ψ is distinct from zero, the point contains

86

something. Of course, this notation, that "something" is denoted by the letter ψ, does not explain anything. But we can now formulate a brief question: What is the physical meaning of the ψ-function?

It is interesting to note that before the answer to this question was given the differential equation for the ψ-function had already been formulated, and physicists had learned how to solve such equations and how to extract physically meaningful results from the solutions. For instance, how to calculate the energy levels in the hydrogen atom (the levels are shown in Fig. 2). It is not difficult to imagine how baffling and exciting the situation seemed to physicists. They were solving problems, and not just ordinary problems, but fundamental ones, dealing with the very basis of our concepts of matter, and did not know what the variable they were operating with was; they did suspect that much more could be extracted from this variable about the properties of electrons and atoms, but did not know how to do it.

We shall begin with a negative statement: the ψ-function is not a field. This means the following: if the ψ-function is distinct from zero at a point in space, this does not mean that a force is applied at this point to some other particle. Unlike the potential of the electric field φ (whose gradient with the opposite sign is the electric field strength) or the product of the electric field strength and the particle's charge, which is the force applied to the particle (this means precisely that φ is a field), there is no simple interpretation of the ψ-function. That the ψ-function is not a field (at any rate, not in the sense

indicated above) becomes apparent from the fact that mathematically ψ is a complex quantity, while all physical quantities are, of course, real quantities.*

The physical meaning of the ψ-function was clarified by Max Born in 1926. He found that the squared modulus of the ψ-function determines the probability of finding a particle at a point in space. Therefore, the wave process we have been discussing so much above describes the probability wave. To be more precise, it describes the wave from which one can determine, by using special techniques, the probability of detecting a particle.

The probabilistic meaning of the wave function (this is the accepted name of the Schrödinger ψ-function) has drastically modified the manner in which events are described in mechanics. The concept of causality has had to be reconsidered.

In classical mechanics the exact calculation of the evolution of a system is made possible by fixing the initial conditions and forces applied to the particles constituting the system. In quantum mechanics the evolution of the system can also be found if the initial state is known. But the state of the system is described in completely different terms. Owing to the uncertainty rela-

---

* This does not mean that complex quantities have no place in physics. They have, and sometimes are much more convenient to use than real quantities. But only real quantities have physical meaning, and if we forgo convenience of calculations, they can always be made to involve only real quantities. In the case of the ψ-function this is impossible. The imaginary part of ψ cannot be eliminated from the mathematical equipment of the theory.

88

tion or, the same thing, to the wave properties of the particles, the state of the system cannot be described in as minute detail as is possible in classical mechanics. The maximum precision within which a mechanical system can be described is given by the wave function. With the wave function known, we can draw conclusions about the results of experiments. However, these results are formulated in probabilistic terms. We cannot say that an event will take place. We can only state that the probability of this event is such and such. This means that quantum laws can be verified only if the experiment is repeated many times (this is always the case in actual conditions).

All this does not mean that quantum mechanics does not provide any unambiguous predictions. It does. For example, an electron in an atom has a quite definite energy. The measurement of the energy of the electron, by means of the ionization potential for example, yields an unambiguous result no matter how often the experiment is repeated. But if we measure the coordinate or momentum of the electron in an atom, we will get different results, whose probabilities are determined by the wave function.

Now a few words about the mathematical tools of quantum mechanics. The study of the motion of particles becomes possible when we know the forces acting on the particles. We know how this is done in classical physics. Newton's equations serve to achieve this purpose. By writing these equations (we do know the forces!) and fixing the initial conditions we can, at least in principle, determine the position of particles at any subsequent moment of time. But all that we have said

earlier (concerning the uncertainty relation and the wave properties of particles) proves conclusively that this method fails when the motion of atomic particles is studied. An essentially new approach is required.

The formulation of basic equations always has something of phenomenology (of description) in it. *Fundamental* equations cannot be derived (at least not until a more general theory is developed). They represent the most compact formulation of the experimental data. Of course, when physicists were obtaining these equations, they were not out in the dark, but by no means did they have a rigorous theoretical scheme ... .

Quantum mechanics has its own rigorously developed, logically consistent mathematical approach to the calculation of physical quantities, an approach that enables us to predict the results of a wide variety of experiments. The motion of atomic particles with velocities small when compared with the speed of light is described by the famous Schrödinger equation, which plays the role that Newton's equations play in classical mechanics. The ψ-function is the solution of this equation. Another mathematical approach (completely different in form) was developed by Werner Heisenberg. Later it was demonstrated that the two methods are identical from the point of view of the physical results obtained.

It will not be possible here to discuss in any detail *how* to describe the motion of atomic particles and what the surprising conclusions one draws from it are. The most surprising is that the conclusions correspond to actually observed

phenomena. The reality is often more fantastic than any fantasy.

Nowadays quantum mechanics permeates all physics. It explains the structure and properties of solid bodies, the emission and absorption of light by atoms, the superfluidity of helium, and the superconductivity of metals. Quantum mechanics is needed to understand the processes inside the stars and to explain the phenomenon of life.

Introduction*
to the Next Five Chapters

Solid State Physics

Different points of view are possible when laymen (and not only laymen) estimate to what degree the phenomena of nature are known and understood; discrepancies are possible even in the interpretation of the word "phenomenon". We shall give here two opposing points of view.

Viewpoint One. "In this world everything is understood", especially if we mean the properties of macroscopic objects, which can be described in complete detail by means of quantum mechanics. The forces applied to the interacting microscopic particles—atoms, ions, and electrons—are known, and the equations describing the motion of the particles are definitely correct. The *only* problems are those of mathematics.

---

* There are people who read introductions after having finished the book. I am one such person. I fear that while writing this Introduction I was unconsciously addressing those readers who have already read the next five chapters. Anyone scared away by the complexities of the Introduction is invited to skip it, finish the book, and then to return to the Introduction. I hope that with this approach the Introduction will appear less complicated and less scholastic.

Input the composition of the object into a sufficiently powerful computer, teach the computer to solve the equations of quantum mechanics, and you are sure to obtain an answer to any question which is *of interest*. But in fact there is nothing to *be interested in*.

Viewpoint Two. "Everything which is interesting and true marks the frontiers of our knowledge, that is, in the theory of elementary particles, in the properties of black holes, and in the mysterious periodicity of pulsars." There are people who think along these lines; sometimes they commit these thoughts to paper, more or less bluntly.

"Of course, the equations of quantum mechanics are well known, some simple problems can indeed be solved, but we do have to admit, though, that not a single really complicated problem can be solved accurately. Now, what can be simpler than solid hydrogen? It consists of the simplest atoms. And for how many years have scientists been trying unsuccessfully to determine its phase diagram? At what pressure will hydrogen become a metal? Will it be an ordinary metal or a superconductor? And if it is a superconductor, what will its critical temperature be? We need a precise answer, please, because it is not cheap to produce pressures of the order of millions of atmospheres; we must know what we are working for and whether the effort is worthwhile. An infinite number of *really interesting* questions remains unanswered, while the theorists concentrate on the neutrino, which is definitely a useless object." That is the point of view of another group of people.

Who is right? I believe that both are wrong. Interesting, important, new, and fundamental

facts can be found in a number of fields of physics: in the physics of elementary particles and in solid state physics. And sometimes these fields are intertwined in a most unforeseen manner, so that a specialist in the theory of superconductivity finds himself in the thick of the events taking place in the physics of elementary particles. Like the material world around us, physics is a single whole. It is divided in a logical manner into domains, some of which may be developed artificially at the expense of others, but then one finds unexpectedly that as a result all the domains, both those that are being developed and those that have been forgotten, are the losers ... .

One has to admit that the physics of the condensed state of matter is indeed based on the already *understood* laws of nature (there is no need to doubt this) and that its place in science is represented by the lower right-hand square in the diagram on p. 49, that is, the square occupied by nonrelativistic quantum mechanics. Nevertheless, this branch of physics remains a source of new concepts. We emphasize the word "concept" in contrast to facts and results that are transferred to technology in a more or less continuous stream. New concepts are so numerous, sometimes they are so out of the ordinary, that a question arises: What is the nature of the appearance of qualitatively new concepts? Or simply: Where do they come from? I believe that they come from the *infinite*. From the infinite number of particles which constitute a condensed body, from the infinite number of combinations of these particles, from the infinite number of ways by which a condensed body can be affected.

The following five chapters are devoted to several problems of the quantum theory of the solid state in its crystalline phase. We believe that these problems are important. Heat capacity, thermal and electrical conductance, the capability of a body to develop a spontaneous magnetic moment are all undoubtedly important properties. Besides, we have selected those problems that enable us to demonstrate the most important cognitive principle of modern macroscopic physics, which we mentioned in passing in the first chapter: to explain a phenomenon means to describe the motion of the "elementary entities" that, as we find, by no means always coincide with the *structural units* of bodies, that is, with molecules, atoms, and ions.

The solid state is a traditional object for quantum mechanical studies. Attempts, and quite successful attempts, to develop solid state theory began almost simultaneously with the formulation of the fundamental principles of quantum mechanics, with the elaboration of the basic equations, and with the development of the calculation techniques. Quantum mechanics explained the existence of dielectrics, semiconductors, and metals; it has found the reason for the strong dependence of the heat capacity and the electrical resistance of metals on the temperature, established the nature of magnetism, and has deciphered the enigma of superconductivity which for so long resisted the efforts of theorists. The achievements of quantum mechanics in solid state physics and in the physics of the condensed state of matter as a whole are innumerable. The whole of modern solid state physics is, in fact, quantum physics.

95

The last decades are characterized by the development of profound understanding of the dynamics of atomic particles in solids. At present we know the motion of electrons and ions in crystals no worse than the motion of electrons in atoms.

Quantum concepts in solid state physics, as we understand it today, began to take shape at the beginning of the twentieth century. One of the basic results of the quantum approach to the study of the properties of crystals was the concept of quasiparticles. It was found that the energy of the excited state of a crystal can be expressed (close to the ground state, though) as the sum

> energy of excited state = energy of ground state +
> + sum of energies of elementary excitations

and that each elementary *excitation* (of the entire crystal) resembles in its properties a quantum particle. Therefore an elementary excitation of a crystal is referred to as a *quasiparticle*, and an ensemble of elementary excitations, as a *gas of quasiparticles*. The formula put in the box now takes the form:

> energy of excited state = energy of ground state +
> + energy of the quasiparticle gas.

The first of these formulas is not quite correct since it does not take into account the interaction between elementary excitations. The second formula makes it possible for us to assume that the gas of quasiparticles is *nearly* ideal, that is, to take into account the interaction between

quasiparticles. Quasiparticles, like real particles, may be fermions or bosons. Examples of both groups of quasiparticles will be found in the chapters to follow, while Chapter 3 describes the "laws of social behaviour" of fermions and bosons.

The quantities describing macroscopic properties of solids are expressed in terms that characterize *individual* quasiparticles: their velocities, the mean free path, etc. While the functional dependence of energy on momentum is the same for all particles*, $\varepsilon = p^2/2m$, where only the mass m characterizes a specific particle, the corresponding formulas for quasiparticles are complicated and take a number of forms. The reason is that particles "live" in a uniform and isotropic space, while quasiparticles "live" in a periodic structure, that is, in a crystal that is in the ground state.

At the risk of repetition, let us list a number of characteristic features of solids, taking a solid as being a physical object consisting of an enormous number of particles whose motion is governed by the laws of quantum mechanics.

(1) Atoms, molecules, and ions are the structural units of solids. This phrase, despite its lack of precision, means that the energy of interaction between structural units is small when compared with the energy that must be spent to destroy a structural unit: to decompose a molecule into atoms or ions, to separate an electron from an atom or ion. At the same time, the energy of interaction *between* particles is not small when

---

* We have finally passed on to the domain of non-relativistic theory.

compared with the energy of their thermal motion. A solid is a system of strongly interacting particles. The possibility of introducing quasiparticles is therefore of principal importance, since it reduces the problem of a system of strongly interacting particles to that of a system of weakly interacting quasiparticles.

(2) According to classical laws, the mean energy of thermal motion of a particle is of the order of $k_B T$, and the internal thermal energy of a body is of the order of $N k_B T$, where N is the number of particles in a body. As temperature decreases, the internal energy of a solid approaches zero faster than a linear function. This is explained by the discrete, quantum character of the energy spectrum of solids. As temperature decreases, separate motions are "frozen out". The greater the energy difference between levels the higher the temperature at which the corresponding motion is "frozen out". For this reason different motions in solids can be, and are, observed at different temperatures.

(3) In real crystals particles are not always arranged in an order which is "forced" on them by the principle of minimum energy. The energy of the crystal increases if an atom or a group of atoms are displaced from an ideal place in the lattice (an atom in an interstitial position, a vacancy, a dislocation, a grain boundary between individual contiguous crystals all these are included in the general term "defect of the crystal lattice"). In principle, the atoms making up a given imperfect crystal could rearrange and form a configuration with lower energy. But this would mean overcoming potential barriers which are high compared with $k_B T$. As a result, imperfect

98

crystals exist, and special measures are needed to grow defect-free or nearly defect-free crystals.

(4) When we have a piece of metal in our hand, we have no doubt about the state of aggregation of this matter. We know that a solid is *hard*. But each solid is composed of a great number of various particles that interact in a different manner with themselves and with other particles. What appears to be a solid is often a *gas* (of quasiparticles*) or a *liquid* (such as that of electrons in metals, which flow around the ionic skeleton of the crystal lattice). Electrons and holes in a semiconductor subjected to a high-frequency field simulate the properties of an electron-positron plasma.

(5) The motions of various atomic particles in a solid are so different that often some of them can be "ignored" when other particles are studied. For instance, the velocity of motion of ions in a solid is so small when compared with that of electrons that there exists a special method of calculation (the so-called adiabatic approximation) in which the ions are assumed to be fixed while the motion of electrons is being considered, the motion of ions being determined by using the characteristics of electrons averaged over their fast motion. A measure of accuracy of the adiabatic approximation is the square root of the ratio of the masses of the electron and ion. It would seem that $\sqrt{m/M}$ is not a very small quantity; however, it is sufficiently small for the zeroth approximation to give a correct picture of the motion of ions and electrons.

* And sometimes of particles as well: the spins of atoms in a paramagnetic at $T \gg T_{\mathrm{C}}$ (T_{C} is the Curie point) form a *gas* of magnetic needles (see Chapter 6).

The motion of atomic particles in solids can be divided into four types.

(a) *Diffusion* of lattice atoms or foreign atoms. Sometimes, as a result of fluctuations, the kinetic energy of particles may be so high that it will exceed the potential barrier separating one unit cell of the crystal from another, and the particle which thus has "gathered its forces" "rolls over" into the neighbouring equilibrium site. Usually the probability of this process, w, at room temperature is extremely low, but increases as we approach the melting point. The duration of the "sedentary" lifetime of the atom is much longer than the time of its displacement, that is, the atom makes only infrequent random jumps. The diffusion coefficient is proportional to w. Diffusion is a rare example of the classical motion of atoms in solids.

(b) Still more infrequent is *quantum diffusion*. It was discovered in solutions of helium isotopes, which exist in the solid state only under high pressures and at temperatures close to absolute zero. In some special cases the tunnelling of atoms from one equilibrium position to another is possible. The possibility of tunnelling delocalizes impurity particles or vacancies, by turning them into very special quasiparticles, the so-called impuritons and vacancions. This transformation is manifested by the diffusion coefficient remaining nonzero at absolute zero temperature.

(c) Quasiparticles are the *collective motions* of particles in a solid, appearing on the *atomic scale*; for example, vibrations of atoms in crystal lattices. Although all atoms take part in each vibration, this vibration is definitely on the atomic scale, since at high temperatures the

mean energy of each vibration (phonon) approximately equals $k_B T$ and is still lower at lower temperatures. Another example of a collective motion is the electron excitation of an atom or a molecule, arising, for example, from the absorption of a light quantum. This motion is not localized in a definite unit cell of the crystal, but "travels" from one cell to another in the form of a *Frenkel exciton**. The mean energy of the exciton is of the same order of magnitude as the energy of the excited state of an individual atom.

The whole world surrounding us is made up of particles of three types: electrons, protons, and neutrons. The diversity in the world stems from that of structures built from these particles. The number of types of quasiparticles is incomparably greater because the motions of atomic particles in solids are infinitely varied. The diversity of types of quasiparticles often overshadows the fact that quasiparticles, as normal particles, may constitute structural elements, that is, building blocks for more complicated entities. Conduction electrons may oscillate. Waves propagate through the electron liquid:

wave $\longrightarrow$ quasiparticle.

In this case the quasiparticle is the plasmon.

Although all electrons have identical charge and therefore are repelled from one another, in a number of metals electrons manage to form a

* The word "exciton" comes from "excited". Owing to the interaction between the atoms of the solid, the *excited* state of an individual atom has been transformed into a quasiparticle.

sort of a molecule, the so-called electron "pair". It was understanding this that led to the explanation of superconductivity, that is, of the phenomenon which for forty-five years remained a tantalizing enigma of solid state physics.

Scientific articles and treatises on solid state physics operate nowadays more and more often with entities made up of quasiparticles: biexcitons, biphonons, light-excitons, and even exciton droplets.

(d) *Superconductivity.* At low temperatures (close to absolute zero) electrons in many metals may participate in a motion which is quantum by nature but macroscopic in scale. This is the superconduction current.

A characteristic feature of the superconduction current is the precise synchronism in the motion of individual electrons; this synchronism results from the interaction of electrons via phonons. This motion is not damped out. With no exaggeration, this motion can *indeed* exist perpetually if, of course, the metal's temperature is kept at a sufficiently low level. The knowledge of the "wonders" of quantum mechanics may blunt our capacity to be surprised: indeed, we already know that the motion of electrons in atoms is perpetual. But it must be added that the nature of the stability of the superconduction current is not similar to the stability of the motion of electrons in an atom.

An electric current is *not a condition* for the existence of the superconducting state. Furthermore, the energy of a superconductor with a current is higher than that without a current. Hence, a superconductor with nonzero current is in a metastable state, its stability being ensured

102

by the fact that to abandon the "collective" of electrons moving synchronously an electron has first to break away from the "pair"; in other words, it has to overcome a certain energy barrier (energy gap). This, however, is not allowed by the laws of conservation of energy and momentum (provided that the superconduction current is not too high; a high current destroys superconductivity). Unfortunately, a detailed description of superconductivity—one of the most exciting quantum phenomena—will not be found in this book. The only argument that the author can offer for justifying this is that superconductivity deserves a book all its own.

(6) The knowledge of what quasiparticles are responsible for is characteristic of our understanding of the energy structure of solids. Experiments rarely lead to misunderstandings. This does not mean that the explanation of each experiment will be immediately clear. Far from it. Often it takes many years for experimental results to be given a complete and final explanation. However, it is very rare for physicists not to be able to pinpoint the responsible party. For some phenomena all quasiparticles are responsible. For instance, heat is transferred by phonons, electrons, and magnons, while the electric charge is transferred only by electrons; the quasiparticles responsible for the temperature dependence of the magnetic moment are magnons, that is, quanta of oscillation of atomic spins.

(7) In most cases all degrees of freedom of the atomic particles in a solid can be divided at each temperature into two groups. Despite the usual meaning of the word, the degrees of *freedom* are interrelated by the forces of interaction. For

example, the spin of one atom interacts with that of another. The energy of interaction, U_{int}, is small when compared with $k_B T$ for some degrees of freedom, and great for other degrees of freedom. As always, here the product $k_B T$ is a measure of the energy of thermal motion. If $U_{int} \ll k_B T$, then the corresponding degrees of freedom behave as an ensemble of *particles* of a weakly non-ideal gas, and if $U_{int} \gg k_B T$, then the corresponding degrees of freedom are, as a rule, ordered, and their motion can be described, also as a rule, by introducing *quasiparticles.*

The above paragraph is of critical importance. It demonstrates that in a predominant number of cases the properties of solids, that is, systems of strongly interacting particles can be described as the properties of *gases* consisting of particles or quasiparticles. Let us emphasize that this is not merely a convenient technique. The very formulation of the problem is to explain macroscopic properties by means of something *elementary*; this requires that the "something" be microscopic. Otherwise we find ourselves in a vicious circle.

Nevertheless, sometimes the situation is much more complicated. The point is that at certain temperatures for specific materials the properties do not change smoothly but in steps. This step change in properties is called the *phase transition.* Phase transitions are observed not only in solids. The transition of liquid to vapour, that is, boiling, is an example of a phase transition. Or the transition of liquid to solid, that is, crystallization. However, these transitions are also possible within a single state of aggregation.

Iron is transformed at $T = 910\,°C$ from α-

phase into γ-phase, and at $T = 1400\,°C$ into δ-phase (these phases differ in the arrangement of atoms within a unit cell and its dimensions). Phase transitions in which some specific property suddenly appears are possibly the most interesting: the spontaneous magnetic moment that appears in the transition from the paramagnetic state to the ferromagnetic state, the ability to conduct the electric current at zero resistance that appears in the superconducting state, etc.

The theory of phase transitions is one of the chapters in thermodynamics and statistical physics, a chapter which contains many interesting results and many as yet unsolved problems. A clear and all-embracing classification of phase transitions is available. Transitions of the first and second order are distinguished. Boiling and crystallization are first-order phase transitions; transitions to the ferromagnetic or superconducting state are second-order phase transitions. Sometimes transitions from one crystal phase to another are of the first and sometimes of the second order.

In first-order transitions, the phases to the left and right of the transition point are usually very dissimilar. One phase replaces another simply because it is more advantageous from the energy point of view. For the transition to take place, it is necessary that the potential barrier separating the phases be overcome. As a result, supercooling and superheating are possible.

In second-order transitions, the situation is different. At the transition point (at $T = T_{\mathrm{C}}$) one of the phases "peters out" and is replaced by a different phase. The phases are indistinguishable at $T = T_{\mathrm{C}}$. But the difference between the

phases is the greater the further we move away from the transition point. Thus, as temperature is lowered from the Curie point, T_C, the magnetic moment of a ferromagnetic increases, and at $T = T_C$ and above T_C it is zero.* Likewise, as we move to lower temperatures away from the temperature of superconducting transition, the number of "paired" electrons, that is, electrons moving without resistance, rises.

It is impossible here to discuss the behaviour of systems in second-order phase transitions, although this topic is exciting to many a scientist. We shall emphasize the following: close to the point of second-order phase transition a solid behaves in such a manner that no artificial technique could be used to describe its behaviour in terms of a gas (either of particles or quasiparticles). Atomic particles located at large distances from one another are found to interact, and this interaction is of a specific, nonforce nature. The probability of collective motions is just as high as that of individual motions.** The anomalous increase in the role of collective motions is confirmed by experiment: specific heat increases and magnetic susceptibility tends to infinity. One additional remark: a solid in the vicinity

---

* This will be discussed in more detail in Chapter 6.
** I clearly understand how difficult it must be for a nonspecialist to grasp these last sentences. The only help I can offer is a comparison between one incomprehensible thing and another one, just as esoteric. The nonforce interaction close to the point of phase transition resembles the quantum nonforce interaction which results in the Pauli exclusion principle. However, in the case under discussion, the interaction is not of a specific quantum nature—second-order phase transitions can occur in classical systems as well.

of a second-order phase transition acts as a *blob of matter*, that is, as a system which is not subdivided into components. Hence the interest in this field displayed by physicists from "another world", namely, specialists who try to decipher the "construction" of *elementary particles*. This interest has been spurred on by the sudden similarity in the mathematical techniques of the two theories.

The theory of second-order phase transitions is not the only point of contact between the physics of macrocosm and that of microcosm, that is, the physics of elementary particles. Attempts have been made in the physics of elementary particles to base the theory of these particles on the properties of space-time. Space-time is a vacuum as far as particles are concerned, the ground state of the material world. With such an approach, particles are elementary excitations of the "vacuum".

This standpoint is very similar to the approach widely used in solid state physics of introducing quasiparticles and studying their properties. The "vacuum" with respect to quasiparticles is the crystal in its unexcited, i.e. ground, state. Quasiparticles are excited states of crystals. An enormous wealth of data has been accumulated in this field. These data enable us to relate the properties of the ground state of crystals of different types ("vacuum") to those of the elementary excitations (quasiparticles) in these crystals. This knowledge may prove useful in understanding the relationship between the properties of space-time and those of particles. Indeed, quite a few of the results obtained in solid state theory are based only on the most general arguments.

107

This is not as fantastic as it might seem: the first publications "bridging the gap" between these "physics", so far removed from each other, are now appearing in the literature.

* * *

The main chapters of the book still lie ahead of the reader, the chapters in which I attempt to describe the specific properties of solids, using almost no formulas. Some of the readers may ultimately consider these details dull compared with the general discussion. This cannot be helped, however: no cognition is possible without details.

Chapter 2

Phonons

Structure of crystals

This small chapter is equivalent to the description of the local geography in an adventure novel, or of the landscape in a lyrical story. The "life" of quasiparticles that we are going to discuss proceeds within a crystal in a comparatively simply organized system of atoms or ions. It is sufficient to know the arrangement of the atoms which make up a *unit cell* in order to be able to "build" the whole crystal by simple translation of this cell.

The crystallization of atoms, that is, the transformation of an unordered ensemble of atoms (the ensemble is gaseous or liquid) into a crystal, is definitely not an example of divine predestination but simply a tendency for a system to lower its energy as the temperature is lowered. Therefore, it is natural that the crystal structure is determined by the forces acting between atomic particles. The nature of these forces was explained by quantum mechanics; thus quantum mechanics explained the structure of solids. Obviously, this was an explanation in principle. It explained why the dimensions of a unit cell are approx-

$\odot$ ions Na $\bullet$ ions Cl $\odot$ ions Cs $\bullet$ ions Cl

(a)   (b)

Fig. 5

imately several angstroms and not, say, several millimetres.

The actual identification of crystal structures is a complicated experimental problem which often (if the unit cell is complex) requires the application of various techniques (the most widely used of which is the diffraction of X-rays) and of rather bizarre mathematical methods. Nevertheless, most crystal structures are now well known. A handbook can tell us about the structure of the unit cell of common salt, the length of the cube edge, and the spacing between the ions Na^+ and Cl^- (Fig. 5a).

The periodical arrangement of atoms in space makes possible a strict classification based on listing the types of symmetry of crystals. The symmetry of a crystal is a set of operations which, being mentally applied to a crystal, transform it into itself. These are rotations about axes,

reflections in planes, helical axes, glide-reflection planes, and translations. It is interesting to note that a combination of "rotational" elements of symmetry (axes and reflections) and translations, without which no periodic structures are possible, results in very special restrictions.

For example, a crystal can possess only axes of 2-, 3-, 4-, and 6-fold symmetry. But nature has no crystals whose lattices would transform into themselves when rotated through an angle of $360°/5 = 72°$. All possible types of symmetry in crystal lattices* were determined at the end of the nineteenth century. Their number, as has been established by the Russian crystallographer E. S. Fedorov, is quite great: 230.

We need to be aware of several crystallographic terms from now on; the first is the *unit cell*, by a periodic translation of which the whole crystal is built. The unit cell is determined by three vectors (often referred to as basis vectors): $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ (Fig. 6). The unit cell volume is

$$v_0 = \mathbf{a}_1 \, [\mathbf{a}_2 \mathbf{a}_3].$$

If all three vectors are mutually perpendicular, then $v_0 = a_1 a_2 a_3$, and if in addition all three have identical magnitude, $a_1 = a_2 = a_3 = a$, then $v_0 = a^3$. The last of these formulas will be used throughout the book for evaluations. The density of the crystal is $\rho = M/v_0$, where M is the mass of atoms within a unit cell of the crystal. It is very important to remember that the linear size of a unit cell is close to the atomic diameter: the simplest model of a crystal is of

---

* A crystal lattice is an imaginary construction: a system of straight lines connecting the atoms of the ideal crystal.
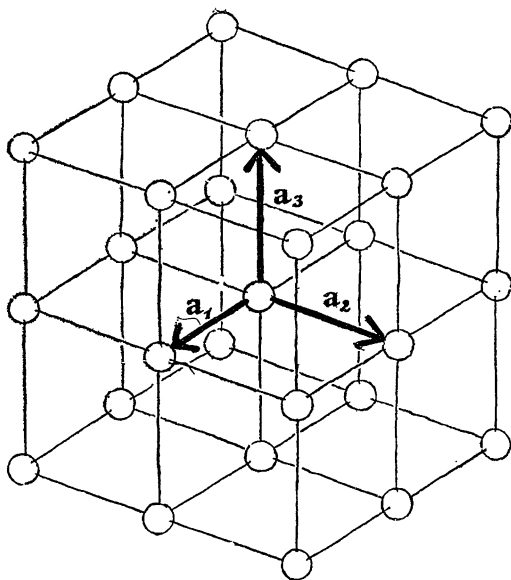
Fig. 6

spheres closely packed into a box. This model gives a fair representation of the crystal structure. There are lattices which are referred to in just this manner—as close-packed lattices. Solid state physics often uses a convenient though arbitrary quantity: atomic radius.

Reciprocal space. There is a formal method enabling us to construct a new crystal lattice, the so-called reciprocal lattice, out of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$. This is done by means of basis vectors

$$\mathbf{b}_1 = \frac{[\mathbf{a}_2\mathbf{a}_3]}{v_0}, \qquad \mathbf{b}_2 = \frac{[\mathbf{a}_3\mathbf{a}_1]}{v_0}, \qquad \mathbf{b}_3 = \frac{[\mathbf{a}_1\mathbf{a}_2]}{v_0}.$$

(23)

(a)   (b)
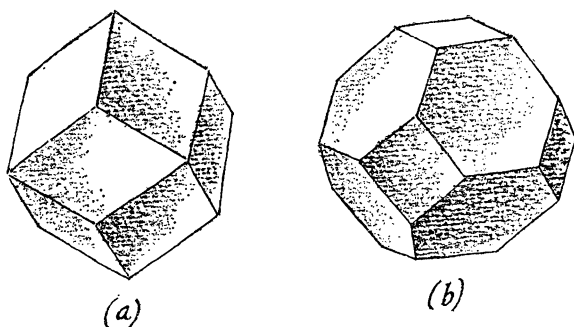
Fig. 7

The unit cell volume of this lattice is $b_1[b_2 b_3] = 1/v_0$. To prove this equality is a good exercise in vector algebra. The spacing between the points of the reciprocal lattice is measured in inverse centimetres. Hence, the lattice is called a reciprocal lattice. By using all the symmetry elements of the crystal, we can build a new unit cell of the same volume from the unit cell of the reciprocal lattice, but one that has all the symmetry elements of the crystal. It is called the *first Brillouin zone*. Figure 7 shows the first Brillouin zone for the body-centred (a) and face-centred (b) cubic lattices. The whole reciprocal space can be filled, with no "gaps", both by the unit cell and the first Brillouin zone. Quasiparticles "live" against the background of the boundaries of Brillouin zones.

Atoms as microscopic pendulums

The motion of atoms and ions with respect to their equilibrium positions in the lattice resembles small-amplitude oscillations of a pendulum.

It should be kept in mind, however, that this statement is not quite exact. Sometimes the amplitude increases so much that an atom leaves its equilibrium position and jumps either into the space between atoms (in this case we speak about interstitial atoms, or interstitials) or into a neighbouring lattice site that happens to be free at the time. At room temperature (as well as at lower temperatures) such jumps are very rare, but their frequency increases as we move closer to the melting point of the crystal.

Atomic motions in the lattice are perpetual, and the range of these motions (physicists call this characteristic the amplitude of vibration) is the greater the higher the temperature. However, the most essential difference between an atom vibrating about its equilibrium position and an ordinary "classical" pendulum is the fact that the atom is a "quantum" pendulum (oscillator), and a quantum oscillator cannot be stopped.

Atoms keep moving even at absolute zero of temperature. This specific quantum form of motion of the atom (it was mentioned in Chapter 1, see p. 39) is called *zero-point vibrations*. The amplitude of zero-point vibrations is usually very small (much smaller than the interatomic spacing), but there does exist one system—namely, solid helium—in which the amplitude of zero-point vibrations exceeds the average distance between helium atoms. As a result, helium never freezes and remains liquid even at absolute zero. Helium can be turned into a crystal only by applying a pressure of about 26 atm.

Helium is a quantum liquid. It resembles a classical liquid only in that it can flow as an ordinary liquid. Helium is not the only quantum

liquid in nature. We shall see in the next chapter that electrons in a metal also form a quantum liquid. It seems very probable that the inner layers of a number of stars discovered in recent years also are in a state closely resembling a quantum liquid. If this is true, quantum liquids are not such a rarity in nature.

The impossibility of crystallizing helium at zero temperature is an *extreme* manifestation of the quantum properties that completely destroy the atomic ordering that classical physics imposes at $T = 0$. A less drastic but no less striking violation of the classical ordering is also possible (and has been observed). It is connected with random defects in the "packing" of atoms unavoidable in any crystal: a lattice site which should contain an atom is vacant.

A vacant lattice site is called a *vacancy*. If the vacant site is occupied by a neighbouring atom leaving its own site, the vacancy moves to the vacated site. Vacancies play a very important role in the diffusion and self-diffusion of atoms in solids. For a vacancy to move, the atom must cross the potential barrier (Fig. 8). Therefore these processes are rare at low temperatures, and the diffusion coefficient rapidly tends to zero as the temperature drops. But in some crystals the picture is different. If a crystal is composed of light atoms, these atoms can "squeeze" from one lattice site to another under the barrier; the accepted expression is: by *tunnelling**. Diffusion

* The tunnelling of quantum particles through regions where the classical kinetic energy of a particle is *lower* than its total energy is a direct corollary of the wave properties of particles. It is similar to the partial transmission of light through a translucent plate.

Fig. 8

and self-diffusion do not stop in such crystals even at $T = 0$, and the crystal possesses singular properties: the nodes of the crystal lattice stand out (on average, the arrangement of atoms is periodic), but the atoms are not completely localized in their proper unit cells: they "wander" through the crystal even at $T = 0$! Such crystals are called *quantum crystals*. Therefore we find that in addition to the crystalline, liquid, gaseous, and plasma states of matter, there are two more states of aggregation: quantum liquid and quantum crystal.

Pendulum? No. Waves? Yes

However conditional the analogy between atoms and pendulums may be, let us turn to it again. Let us imagine a system of pendulums, or rather a system of balls connected by springs (Fig. 9). Once we displace one of the balls, it moves its neighbour, which then shifts the next ball, and so on. A wave moves through the system of balls and springs. These waves of motion of the balls are a good illustration of the motion of atoms in solids. Such waves in solids represent *sound waves*.

Like any other waves, sound waves are characterized by wavelength λ and frequency ω. The waves that are able to travel in crystals cannot

116
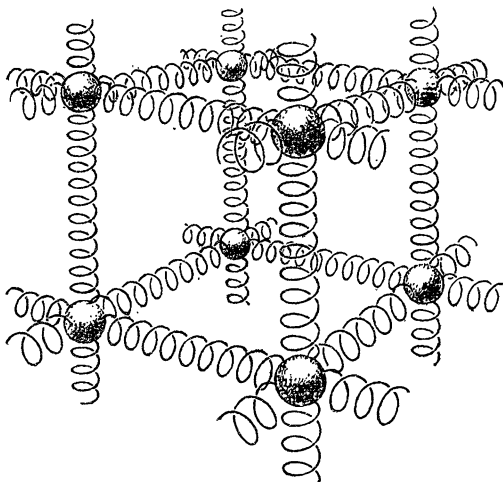
Fig. 9

have arbitrary λ and ω, but must satisfy a specific relation between frequency and wavelength (or frequency and wave vector):

$$\omega = \omega (\mathbf{k}). \qquad (24)$$

This relation is called the dispersion law.

If the forces between atoms are known, then the dispersion law of sound waves can be calculated. However, such a calculation is comparatively cumbersome and can be done only by means of computers. But even if high-speed computers are used, the problem has to be simplified as much as possible. One has to assume, for example, that atoms that are sufficiently far apart do not interact at all.

The dispersion law of sound vibrations given in eq. (24) is the main characteristic of the motion of atoms in a crystal. Knowledge of the dispersion law enables us to calculate a number of properties of crystals (heat capacity, thermal expansion coefficient, heat conduction, etc.). It is not surprising that both theorists and experimenters are interested in the dispersion law of sound waves in crystals ... .

Fortunately, many properties of the dispersion law can be found on the basis of general arguments, without specific calculations. "Fortunately", because these properties do not depend on the approximations which are invariably used in specific calculations. On the contrary, they can serve, and do serve, as a means of checking the numerical solution of problems involving the vibrations of atoms in a specific crystal.

Naturally, in this context we cannot *derive* the general properties of the dispersion law. The presentation will thus be limited to a few statements.

(1) The number of *types* of different waves that can propagate in a crystal is $3r$, where r is the number of different types of atoms (or ions) in a unit cell of the crystal. Thus, six types of waves can propagate in common salt, and only three in solid argon, since the NaCl unit cell contains two ions, while the Ar unit cell contains a single argon atom.

(2) Out of $3r$ types of waves three are called *acoustic modes*. Their specific feature is that for small wave vectors (i.e. with long wavelengths) their frequencies are low and tend to zero as the wave vector **k** vanishes. The remaining $3r - 3$

wave modes are called *optical modes*. The term "optical" is rather accidental. The point is that they were first detected by optical methods.

(3) Each of the $3r$ functions $\omega_j(\mathbf{k})$ (the subscript j enumerates the wave modes: $j = 1, 2, \ldots \ldots, 3r$) is a periodic function of its argument:

$$\omega_j(\mathbf{k} + 2\pi\mathbf{b}) = \omega_j(\mathbf{k}),$$
$$\mathbf{b} = n_1\mathbf{b}_1 + n_2\mathbf{b}_2 + n_3\mathbf{b}_3, \tag{25}$$

where n_1, n_2, n_3 are integers, and vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ are basis vectors of the "reciprocal" space which we were discussing at the beginning of this chapter (see eqs. (23)). The periodicity of the frequencies of vibrations is a manifestation of the general property of crystals: the periodic arrangement of atoms in real space "generates" a *periodic* reciprocal $\mathbf{k}$-space (note that the dimension of $\mathbf{k}$ is $[k] = \mathrm{cm}^{-1}$) and enables us to limit the analysis to a unit cell in the reciprocal space, or, more conveniently, to the first Brillouin zone.

The reader may be surprised by these three paragraphs. Why are sound waves (it is indeed sound waves we are talking about) so complicated? We are used to the idea that sound is something quite simple. *Sound* waves in a gas (for example, in air) or in a liquid are oscillations of density or of pressure. It is true that in addition to oscillations of density, shear oscillations can also propagate in solids. In the case of oscillations of density the atoms vibrate along the wave vector $\mathbf{k}$ (longitudinal waves), and in shear waves they vibrate in a plane perpendicular to the wave vector (transverse waves). Two transverse waves are possible in a crystal.

Therefore, three sound waves can propagate in a solid: one longitudinal wave and two transverse

waves. The dispersion law of sound waves is simple. For longitudinal waves

$$\omega = s_l k, \qquad (26)$$

and for transverse waves

$$\omega = s_t k \qquad (27)$$

and the velocity of longitudinal sound waves s_l is always higher than that of transverse waves s_t.

Now, what is the relationship of these "elementary" sound waves to the vibration of atoms in a crystal lattice? The simplicity of the dispersion law of "elementary" sound waves (eqs. (26) and (27)) and of the whole pattern of the propagation of sound is caused by the fact that, in describing the phenomenon, we ignored the atomic (discrete) structure of the crystal and took it as continuous* ... . This approximation (naturally, this approach is only approximate because in reality solids consist of atoms and molecules) is valid only when the wavelength of the vibration is long when compared with the interatomic spacings, $\lambda \gg a$, meaning that we can ignore the atomic structure. But since $\lambda = 2\pi/k$, the condition of applicability of the approximate (macroscopic) approach takes the form

$$ak \ll 1. \qquad (28)$$

The microscopic approach is more accurate and, of course, the conclusions of the macroscopic

---

* A special section of mechanics which studies the dynamic and static properties of solids and liquids in which their atomic structure can be ignored is called *continuum mechanics*. Hydrodynamics and the theory of elasticity are parts of continuum mechanics.

analysis must follow from the more rigorous microscopic analysis. Reread statement (2) above. Under condition (28), the acoustic modes of atomic vibrations transform into ordinary sound waves. When a "simple" sound wave propagates through a crystal, unit cells of the crystal are displaced without deformation with respect to one another. Optical $3r - 3$ vibrations (if they are present, see statement (1) above) describe the vibrations of atoms with respect to one another within each unit cell of the crystal. As a rule, the frequencies of these vibrations are high, of the order of infrared optical frequencies.

Let us return now to our initial model. We compared the motion of atoms in a crystal to the oscillations of pendulums and then decided that we have to reject this analogy. Atoms are not isolated from one another, so that atomic vibrations are not independent, but take the form of waves in the crystal. But each wave (with a specific wave vector $\mathbf{k}$) has a definite frequency ω. Is it justified then to treat the whole wave as a "pendulum", i.e. an oscillator vibrating at frequency $\omega(\mathbf{k})$? A mathematical analysis of the problem shows that this is a very fruitful analogy.

An oscillator with frequency $\omega_j(\mathbf{k})$ is an elementary form of atomic motion in crystals, a structural unit of the vibrational motion of atoms in crystals. And although the motion of each oscillator is elementary, that is, cannot be decomposed, all atoms of the solid participate in this motion. Each oscillator moves independently of all others. Hence, the energy of the vibrational motion of atoms is the sum of the energies of individual oscillators.

This is a most important conclusion. The energy of *interacting* atoms (ions or molecules) is thus found to be the sum of the energies of *noninteracting* oscillators. We see that an oscillator plays the part of a structural unit of motion in crystals. A structural unit of matter in crystals is, of course, an atom, a molecule, or an ion (see in Chapter 1 the section What are the components of the surrounding world?).

If the reader is not afraid of seemingly too bold analogies, he will accept the statement

> a solid is a "gas" of oscillators

—of course, only from the "point of view" of the vibrational energy of the crystal, since it is precisely in the gas that the energy is the sum of the energies of individual particles.

The dispersion laws characterize the dynamic properties of crystals in great detail. Sometimes, however, the description is too detailed. Often it is sufficient to know how many vibrations there are per interval of frequencies $\Delta \omega$. In order to determine this function (the so-called density of states, denoted in this book by $\nu (\omega)$) we have to calculate how many wave vectors $\mathbf{k}$ there are between two isofrequency surfaces:

$$\omega_j (\mathbf{k}) = \omega \quad \text{and} \quad \omega_j (\mathbf{k}) = \omega + \Delta \omega$$

and to take the sum over all j (from 1 to $3r$). Function $\nu (\omega)$ is very complicated (for instance, see Fig. 10). Its determination is one of the practical problems of the modern dynamic theory of crystals and involves the use of computers.

However, here we are not interested in the admittedly important details of the dependence of
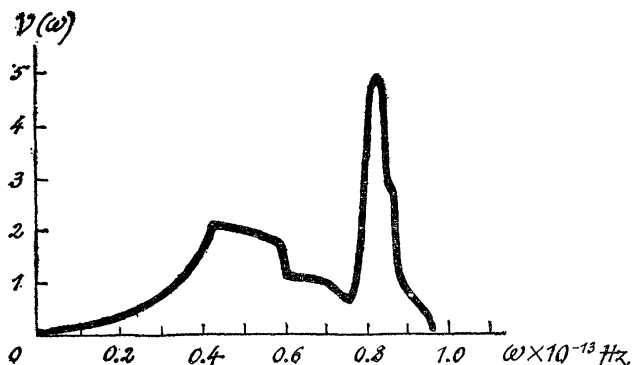
Fig. 10

the density of states on frequency; we want only the most salient features of function $\nu(\omega)$. In order to determine these features we limit the analysis to the so-called Debye approximation, that is, we assume that for $k < k_D$ the sound waves are identical to the waves in the continuum (see eqs. (26) and (27)) and that there are no waves for $k > k_D$. The limiting value k_D can be found from the natural requirement that the total number of vibrations be equal in crude theory and in exact theory. The result thus obtained is also natural: $k_D \approx \pi/a$, that is, no waves with wavelengths smaller than the spacing between the neighbouring atoms can propagate in the crystal.

The density of states in the Debye approximation has a very simple meaning. In the range from the lowest frequency to a certain maximum frequency, $\nu(\omega)$ is proportional to ω^2, and above

123

this upper limit the density is, of course, zero:

$$\nu(\omega) = \begin{cases} \dfrac{3V}{2\pi^2 \tilde{s}^3}\,\omega^2, & \omega < \omega_{\max}, \\ 0, & \omega > \omega_{\max}. \end{cases} \quad (29)$$

Here V is the volume of the crystal, and $\tilde{s}$ is the "mean" velocity of sound: $3/\tilde{s}^3 = 2/s_t^3 + 1/s_l^3$. The maximum frequency (referred to as the *Debye frequency*) is equal to $\omega_D = (6\pi^2)^{1/3}\tilde{s}/a$ if the unit cell volume is assumed to be equal to a^3. The total number of oscillators must be $3N$, where N is the number of unit cells in the crystal, that is, $V = Na^3$. As the speed of sound in solids is of the order of 10^5 cm/s and the lattice parameter $\approx 3 \times 10^{-8}$ cm, then $\omega_D \approx 10^{-13}$ s^{-1}. Although the true density of states is very far from the Debye density (see Fig. 10), the estimate of the maximum frequency is realistic; it is at least of the right order of magnitude.

The replacement of the true density of states $\nu(\omega)$ by the Debye density is perfectly acceptable if the crystal is comparatively simple and the unit cell contains only one atom. If, however, the crystal is more complex, then the fit of the experimental data to the curve can be achieved only if we assume that in addition to oscillators with frequencies $0 < \omega < \omega_D$ there are also oscillators with discrete frequencies. The number of these oscillators is $(3r - 3)N$, where r is, as before, the number of atoms per unit cell, and N is the number of unit cells in the crystal.

To conclude this section, we want to emphasize that the picture presented above is a corollary of the application of classical Newtonian mechanics to a system composed of a large number of

124

atoms. Strictly speaking, the number of atoms is infinite since we have not taken into account the boundaries of the crystal.

Quantum oscillator

From the standpoint of classical mechanics, the energy of an oscillator can be arbitrary: the larger the range of the oscillation (its amplitude) the higher its energy. If the maximum deflection of the oscillator from its equilibrium position is denoted by the symbol u, and its energy by E, then

$$E = \frac{m\omega^2 u^2}{2}, \tag{30}$$

where m is the mass of the oscillator, and ω is its frequency.

Quantum mechanics states: the energy of an oscillator can assume only discrete values:

$$E = E_n = \hbar\omega \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, 3, \ldots . \tag{31}$$

These levels are shown in Fig. 11.

Do we always have to take into account the quantum properties of an oscillator? Let us begin with a standard pendulum, oscillating at the frequency of a clock pendulum ($\omega = 2\pi \text{ s}^{-1}$). The allowed energy levels of this pendulum are separated by intervals that are approximately equal to 6×10^{-27} erg. Let us calculate, for the benefit of a reader who does not know whether this is a lot or a little, the amplitude of oscillations of a pendulum as given by eq. (30) if the energy is 6×10^{-27} erg.

125
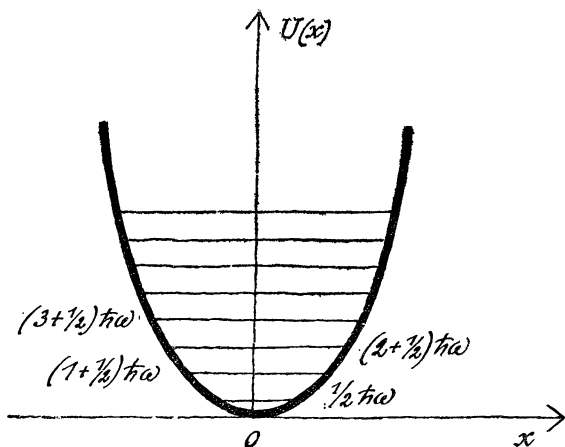
Fig. 11

Even for a mass of 1 g ($m = 1$ g) the amplitude is fantastically small: less than 10^{-13} cm (i.e. less than the diameter of an electron, less than the diameter of an atom by a factor of a hundred thousand). It is clear that it would be meaningless to take into account the discreteness of the energy spectrum when the motion of macroscopic bodies is studied: the allowed energy levels are so close to one another that the gaps can simply be ignored.

The mean energy of a particle in a gas of classical particles* is $(1/2) k_{B}T$, where T is the gas temperature. A solid is a "gas" of oscillators. However, the energy of the oscillator equals the

---

    * Gases of quantum particles are discussed in Chapter 3.

sum of its kinetic energy and its potential energy. On average, these energies are equal. The mean energy of each oscillator equals, at least according to the laws of classical physics, $k_B T$. This enables us to determine when the formulas of classical mechanics are adequate and when quantum mechanics has to be used. To achieve this we have to compare the spacing between the energy levels of the oscillator, $\hbar\omega$, with the energy, $k_B T$.

First, at any temperature there are oscillators for which $\hbar\omega \ll k_B T$ and to which, therefore, classical mechanics is applicable. Second, classical mechanics is applicable *to all oscillators* only at $k_B T \gg \hbar\omega_D$ and, hence, it can be applied to the study of properties of solids with no risk of making a mistake.

The quantity $\hbar\omega_D$ is called the *Debye energy*, and if it is divided by the Boltzmann constant k_B, it is called the *Debye temperature*, denoted, as a rule, by Θ:

$$\Theta = \frac{\hbar\omega_D}{k_B}. \tag{32}$$

In most solids $\Theta \approx 100$ K, that is, it is several times lower than room temperature (20 °C = = 293 K). As a result, almost all solids show no quantum peculiarities under normal conditions. However, there are elements, such as, for example, beryllium, in which the Debye temperature is anomalously high ($\Theta_{Be} \approx 1000$ K). The Debye temperature is one of the most important characteristics of crystals. Each handbook of physical constants, if it is at all detailed, lists the Debye temperatures of various solids.

Classical and quantum heat capacities of crystals

It is convenient to begin the calculation of heat capacity by determining the internal thermal energy of a body, $\mathscr{E}_{osc}$*, since by definition heat capacity C is the change in the thermal energy of the body when its temperature is changed by one degree: $C = \partial \mathscr{E}_{osc}/\partial T$. The vibrational energy is that of the gas of oscillators, so that it is natural to calculate the mean energy of this gas and thus find the mean energy of a single oscillator. We have shown above that at temperatures below the Debye temperature a quantum analysis is required. Of course, it can be used at high temperatures as well, because it is more exact. Max Planck has shown that, as follows from eq. (31), the mean energy of an oscillator is

$$\bar{E} = \frac{1}{2} \hbar\omega + \frac{\hbar\omega}{\exp\left(\hbar\omega/k_B T\right) - 1} \, . \qquad (33)$$

Within the classical limit $k_B T \gg \hbar\omega$, we obtain, as could be expected, the correct result $\bar{E} = k_B T$. This is an additional illustration of the validity of the correspondence principle.

In order to find the internal thermal energy of a crystal, the energies of all oscillators have to be added up. This means that $\bar{E}$ must be multiplied by the density of states $\nu(\omega)$ and integrated over the full range of oscillator frequencies. This finally yields

$$\mathscr{E}_{osc} = \mathscr{E}_0 + \int_0^\infty \frac{\nu(\omega)\, \hbar\omega\, d\omega}{\exp\left(\hbar\omega/k_B T\right) - 1} \, . \qquad (34)$$

---

* By denoting this quantity $\mathscr{E}_{osc}$, we emphasize that we mean the vibrational energy of the atoms in the solid.

128

Before going any further, let us briefly remark on (1/2) $\hbar\omega$ in eqs. (31) and (33). We have already mentioned that a quantum pendulum cannot be stopped. Indeed, the minimum oscillator energy, the energy of its zero-point vibrations (at $n = 0$), is equal to (1/2) $\hbar\omega$. The total energy of the zero-point vibrations of the lattice atoms is denoted by $\mathscr{E}_0$. This is the vibrational energy of the crystal at absolute zero temperature. However, the crystal need not necessarily be cooled to absolute zero to detect the zero-point vibrations. There are phenomena in which zero-point vibrations can be observed at a finite temperature as well. For example, the main reason for the dissimilarity in the properties of crystals composed of different isotopes of the same chemical element is the difference in the energy of zero-point vibrations.

Equation (34), which makes it possible to calculate the internal vibrational energy of a solid, demonstrates that all we need to know for this is the density of oscillator states ... . Nothing but that. This is why function $\nu(\omega)$ is considered to be so important. We also have to remind the reader that knowing $\mathscr{E}_{osc}(T)$ enables us to find, in fact, all equilibrium (thermodynamic) characteristics of the crystal*: its heat capacity, compressibility, thermal expansion coefficient, etc.; they are determined without introducing any new assumptions concerning the

---

* To be precise we note that one also has to know the dependence of all these quantities on pressure. This dependence is also determined by the forces of interaction between atoms and is included in the density of states (for example, as a dependence of the Debye temperature on pressure).

character of the motion of atomic particles. This is the realm of a special science that does not resort to models: thermodynamics based on such general principles as the law of energy conservation ... .

The quantum character of the gas of oscillators is manifested in all thermal properties of a solid at low temperatures. The manifestation is clearest in heat capacity as a function of temperature. At high temperatures (we know now that this means $T \gg \Theta$) the quantization of energy is unimportant, and we obtain a classical result

$$\mathscr{E}_{\text{osc}} = 3Nrk_{\text{B}}T \tag{35}$$

and thus heat capacity C is independent of temperature; calculated per mole of matter, it is equal to $3R$ (where R is the gas constant, $R = 1.9858$ cal/mol·K). The equality $C = 3R$ is called the *Dulong-Petit law*. This high rank (that of a law) does not protect it from restrictions. In accordance with eq. (34), at $T \ll \Theta$ the vibrational energy tends to zero with temperature, and heat capacity (again per mole of matter) equals

$$C = \frac{4\pi^4}{15}\,3R \left(\frac{T}{\Theta}\right)^3. \tag{36}$$

This result, derived by Debye and found to be in perfect agreement with experimental data, was one of the triumphs of quantum mechanics in explaining the properties of macroscopic objects.

The main contribution to the vibrational energy of crystals at low temperatures is provided by acoustic waves. The energy of the corresponding oscillators is low, and therefore they are easily

excited. Optical waves (even if these are allowed) are not excited at low temperatures: thermal energy is insufficient for this excitation at $T \ll \Theta$.

Phonons

Now we have to master one of the main concepts of the quantum theory of solids. We shall try to demonstrate how the corpuscle-wave dualism, which was described in sufficient detail in Chapter 1, enables us to develop a convenient, graphically clear language for describing the behaviour of atoms in a solid, a language which is especially convenient at low temperatures.

First, let us construct a scheme of what has been described in the preceding sections of this chapter.

A crystal is an ensemble of interacting atoms arranged in an ordered pattern. Atomic vibrations propagate through the crystal as weakly interacting waves with wave vectors $\mathbf{k}$ and frequencies $\omega_j (\mathbf{k})$. To each wave we can assign an oscillator with frequency $\omega_j (\mathbf{k})$.

But according to Louis de Broglie's ideas, to each wave we can assign a *particle* with energy ε equal to $\hbar \omega_j (\mathbf{k})$ and momentum $\mathbf{p}$ equal to $\hbar \mathbf{k}$. This particle (it is usually referred to as a *quasiparticle*, see Introduction to the Next Five Chapters: Solid State Physics) is called the *phonon*. The phonon is an elementary portion of sound energy with frequency ω, the same way as a photon is an elementary portion of light energy (electromagnetic energy).

The comparison is like this:

wave $\longrightarrow$ quantum oscillator $\longrightarrow$ phonon.

This comparison is based on the fact that the energy of an excited state of each *quantum oscillator* ε_n is equal to an integer number of $\hbar\omega_j$ (**k**):

$$\varepsilon_n = n\hbar\omega_j\,(\mathbf{k}), \qquad n = 1, 2, 3, \dots \, . \tag{37}$$

The number n is thus given a simple and clear meaning: n is the number of phonons of a given type, namely, with momentum $\mathbf{p} = \hbar\mathbf{k}$ and energy $\varepsilon = \hbar\omega_j$, and $1/[\exp(\hbar\omega_j/k_BT) - 1]$ in eq. (33) is none other than the *average* number of phonons with *momentum* $\mathbf{p}$ and *energy* ε at a given temperature. Therefore we introduce a symbol:

$$\overline{n} = \frac{1}{\exp(\hbar\omega/k_BT) - 1} \, . \tag{38}$$

Note that the average number of phonons $N(\varepsilon)$ with a given energy is obtained by multiplying $\overline{n}$ by the density of states $\nu(\varepsilon)$:

$$N(\varepsilon) = \frac{\nu(\varepsilon)}{\exp(\varepsilon/k_BT) - 1} \, . \tag{39}$$

The introduction of the phonon enables us in many cases to describe any solid as a box containing the phonon gas. Phonons, as particles of an ordinary gas, move from wall to wall, colliding with one another.

The phonon gas is the main heat reservoir of a solid. The thermal properties of a solid can only be understood by studying the properties of this unusual gas. And, of course, the phonon gas is a very unusual gas. The most spectacular difference between this gas and an ordinary gas is that the number of phonons in a solid is not constant. The more intensive the thermal motion of atoms, that is, the higher the temperature,

the greater the number of phonons. At high temperatures the number of phonons is proportional to temperature, and as we move closer to absolute zero, the number of phonons also tends to zero, in proportion to the third power of temperature. Both these statements are easily checked if we use the expression that has just been given above for $N(\varepsilon)$, the number of phonons with a given energy.

Like waves, there are acoustic and optical phonons. However, at low temperatures, when the quantum approach is unavoidable, the number of optical phonons is exponentially small (we have already mentioned that thermal energy is not sufficient to create an optical phonon). Besides, not all of the acoustic phonons are excited at low temperatures. Most of the phonons are with energies close to $k_B T$. There is not enough energy for more energetic phonons (the temperature is too low) and the number of low-energy phonons is always small: the density of states is proportional to ε^2 (see formula (29)).

Let us calculate the momentum of these most numerous phonons at $T \ll \Theta$. The transition from k and ω to momentum and energy yields $\varepsilon = sp$ when mostly long-wavelength acoustic phonons are excited, with the dispersion law $\omega = sk$ (s is one of the velocities, either transverse or longitudinal). It is easy to determine from the condition $\varepsilon \approx k_B T$ that the momentum of these phonons is approximately equal to $(\pi\hbar/a) \times \times (T/\Theta)$, since $k_B \Theta \approx \pi\hbar s/a$.

The transition from waves (with wave vectors **k**) to quasiparticles (with momenta **p**) makes it necessary to transfer from reciprocal space (see p. 113) to *momentum space*. The characteristic

size in this "space" (it is convenient to refer to it as "**p**-space") is $\pi\hbar/a$. The momenta higher than this value are simply devoid of meaning — their equivalents can be found in the first Brillouin zone by using the periodicity of the p-space.

We find, therefore, that at $T \ll \Theta$ the thermal phonons occupy a small region close to the centre of the first Brillouin zone. As temperature rises their number increases, and the higher the temperature the more uniform the distribution of phonons in the p-space.

Digression on photons

After this very detailed discussion of phonons, we can return to Chapter 1 and elaborate to a certain extent the statements that were meant to clarify the transition from electromagnetic waves to particles, namely photons.

Formally, the line of reasoning is similar to the one used above. To an electromagnetic wave with wave vector **k** and frequency ω we assign a quantum oscillator with frequency ck, *or* with a particle, i.e. photon with momentum $\hbar\mathbf{k}$ and energy $\hbar\omega$; the average number of photons with momentum **p** is $1/[\exp(\hbar\omega/k_B T) - 1]$; and the mean energy of a photon is $\hbar\omega/[\exp(\hbar\omega/k_B T) - 1]$. The density of electromagnetic energy in a cavity (eq. (19)) is deciphered as follows:

$$E_\omega = \text{density of photon states} \times \\ \times \text{mean energy of photon.}$$

The difference between the density of photon states and $\nu(\omega)$ (compare eq. (19) and formula (29)) is connected, first, with the difference be-

tween the speed of sound and the speed of light, and, second, with the difference in the number of wave modes: a sound wave has one of *three* possible polarizations, and a light wave, one of *two*.

Phonons transfer heat

It is well known that every solid body is a conductor of heat. It is especially easy to explain what heat conduction is if we consider a crystal as a box containing phonons. Let one end of a solid be kept at temperature T_1, and the second end, at temperature T_2, and let $T_2 > T_1$. This means that the phonon's density at the left-hand end of the body (where the temperature is T_1) is lower than that at the right-hand end (where the temperature is T_2). Obviously, the phonons at the right-hand end will "flow" to the left-hand end, trying to equalize the phonon's density over the whole solid (if we gathered particles of a gas in one part of a vessel, the gas would gradually occupy the whole volume of the vessel). As phonons move from the warm end to the cold end, they transfer heat, which is removed at the cold end if it is maintained at a constant temperature. The phonons arriving at this end simply disappear since their number, as we have mentioned above, is a function of temperature only. At the hot end phonons are created, so that constant temperature can be maintained only by supplying heat to this end.

The concept of thermal motion in a solid realized in the motion of phonons in the phonon gas is convenient not only in describing phenomena (the phenomenon in question is heat conduction)

135

but also in calculations. According to the molecular kinetic theory of gases, the heat conductivity $\varkappa$ is equal to the product (1/3) $Cl\bar{v}$, where C is the heat capacity per unit volume of the gas, l is the mean free path (or simply range) of the gas molecules, and $\bar{v}$ is their mean thermal velocity. We are dealing here with the phonon gas. Limiting the analysis to the Debye approximation for the sake of simplicity, we determine that the velocity of phonons is equal to the speed of sound $(\bar{v} = s)$. We already know the heat capacity of the phonon gas (see eqs. (35) and (36)). In order to determine the temperature dependence of the heat conductivity it is now necessary "only" to find the mean free path of the phonons, l_{ph}. The mean free path of particles is determined by collisions, that is, by factors which limit their free path.

Let us try to analyze the collision of phonons. Are we entitled to speak of collisions when in fact we are talking not about real particles but about phonons, that is, quasiparticles (in other words, entities resembling particles, or imaginary particles)? Crystals always contain some sort of inhomogeneity: impurity atoms, boundaries of crystal grains, dislocations*, and finally the

---

* A crystallite or crystal grain is a small area within a crystal with strictly periodic arrangement of atoms. A single crystal is a single crystallite. Crystals which, at least until very recently, were employed in industry and in domestic appliances, were polycrystals made of an enormous number of crystal grains. Dislocations are linear defects. An example of a dislocation is the trace of an incomplete atomic plane within a crystal. This is a frequently encountered defect of crystal structure, called an edge dislocation.

outer surfaces of the sample. The mean free path of phonons is, of course, limited by the inhomogeneities of the crystal. It is convenient to recall at this stage that the phonon is a wave that is scattered as a wave should and, therefore, can be scattered by an obstacle. The wave properties of phonons explain why the limiting role of impurities is not critical for the mean free path, especially at low temperatures. A wave whose wavelength is much longer than the size of an inhomogeneity* responds only very slightly to an obstacle (this is easily established if one observes waves on the water surface at the seashore).

As the number of inhomogeneities decreases, that is, the quality of the crystal improves, the mean free path will increase. What is the limit of this increase? Is the mean free path in an ideal, "infinite", crystal infinitely long? Or are there some factors specific to an ideal crystal that restrict the mean free path of the phonons?

Yes, such factors do exist! Like the molecules of an ordinary gas, phonons collide with one another, and the laws governing the interaction between phonons can be understood by studying the motion of atomic particles in crystals.

Let us analyze first of all what is meant by the word "collide". There would be no "collisions" if the particles (or quasiparticles) did not interact at all. For example, photons in a vacuum practically do not "notice" one another (with a tremendously high accuracy). To translate the scientific jargon, the word "interaction" means

---

* The size of an inhomogeneity is approximately equal to the unit cell dimension, a, while the phonon wavelength, λ_{ph}, is much longer than a at $T \ll \Theta$: as follows from $p \approx (\pi \hbar / a) (T/\Theta)$, we have $\lambda_{ph} \approx a\Theta/T$.

in rigorous terms that the energy of two particles is not equal to the sum of the energies of these particles but has one, or even several, additional terms. These additional terms are called precisely the *energy of interaction*.

The probability of collision and other characteristics of the collision process are expressed via the energy of interaction. But it seems that we have shown that the energy of the phonon gas is the sum of the energies of the individual phonons. We have emphasized that this fact is very important. Yes, we did show it. But what was the basis for this conclusion? It was based on the representation of atomic motion as *small* vibrations (an analogy to the motion of pendulums). Indeed, if we assume that a displacement of an atom from its equilibrium position causes a restoring force proportional to the displacement, then all we have said above is absolutely true. However, the statement that force is a linear function of displacement is not exact! Hence, the name for this approach—*harmonic approximation*.

The amplitude of atomic vibrations being very small, in many cases the harmonic approximation is sufficient. But not in the case of calculating the mean free path of phonons. *The interaction between phonons is determined by higher-order* (after the linear) *terms in the expansion* of the force in terms of the powers of displacement. These terms are called *anharmonicities*. Of course, anharmonicity is important not only in calculating the mean free path of phonons. Were these terms not taken into account, one could not understand the thermal expansion of solids. But this problem will not be discussed here,

The collisions of phonons are very unlike those of ordinary particles. Let us represent a particle in a drawing by an arrow, and the fact of a collision by a circle. Then a collision of two particles can be shown by the diagram
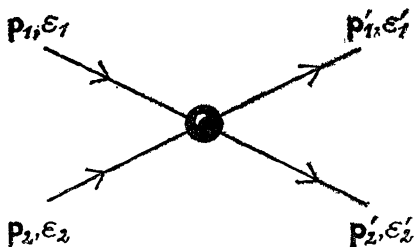
$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_1' + \mathbf{p}_2',$$
$$\varepsilon_1 + \varepsilon_2 = \varepsilon_1' + \varepsilon_2'.$$

The above equations reflect the momentum and energy conservation; obviously, these laws hold when particles of a gas collide.

And here is a drawing that shows one of the possible types of collisions between phonons:

$$\hbar\mathbf{k}_1 + \hbar\mathbf{k}_2 = \hbar\mathbf{k}' + \hbar\mathbf{K},$$
$$\hbar\omega_1 + \hbar\omega_2 = \hbar\omega'.$$

139
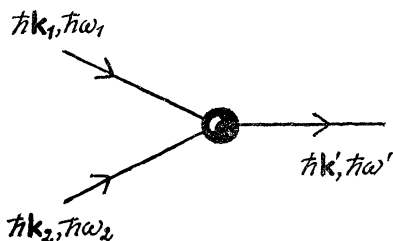
What is striking in the drawing is that the number of particles in a collision is conserved, while the number of phonons is not. But what seems to be the main difference is that the collisions of phonons do not conserve momentum. The crystal lattice in which phonons move also takes part in collisions, by consuming a "part" of momentum. This part, however, is strictly prescribed: $\hbar\mathbf{K} = 2\pi\hbar\mathbf{b}$. The crystal lattice does not always participate in the collisions of phonons. Collisions conserving momentum are called *normal collisions*. Collisions which do not conserve momentum are called *Umklapp processes*. The reason for Umklapp processes lies, of course, in the periodicity of the p-space.

If the momentum of the colliding particles leaves the first Brillouin zone, it has to be "returned". This is the meaning of an "Umklapp process". Umklapp processes offer additional channels of scattering: the lattice is the scene for collisions which are forbidden by the conservation law in a vacuum.

It can be easily understood that only those collisions that are accompanied by Umklapp processes limit the mean free path of phonons. Indeed, if the laws of momentum and energy conservation are satisfied, then the energy and momentum are only redistributed in a collision; this does not slow down the phonon gas which simply flows as a single unit. This gas can transfer any amount of heat, the higher the gas flow velocity the greater the amount.

For an Umklapp process to take place, the collision process must involve at least one phonon with high momentum close to $\pi\hbar/a$. The number of such phonons is great at high temperatures,

and the mean free path included in the expression for heat conduction is very short ($l_{\text{ph}} \approx \approx 10^{-6}$ cm).

However, as the temperature falls, the number of high-momentum phonons becomes smaller and smaller. Remember? At $T \ll \Theta$ almost all phonons are close to the centre of the first Brillouin zone. Phonons with momenta close to $\pi\hbar/a$ have an energy of the order of $\hbar\omega_{\text{max}} = = k_{\text{B}}\Theta$. And this means that their number at $T \ll \Theta$ is exponentially small, that is, proportional to $\exp\left(-\Theta/T\right)$. Naturally, the probability of a collision with such phonons is reduced, so that the mean free path l_{ph} increases (exponentially). By combining the dependence of heat capacity and mean free path on temperature, we obtain

$$\varkappa \sim T^3 \exp\left(\Theta/T\right). \tag{40}$$

As we move closer to absolute zero, the mean free path becomes so large that the finite dimensions of the specimen can no longer be ignored: crystal boundaries are an obstacle to the propagation of phonons, and $l_{\text{ph}} \approx L$, where L is the smallest of the crystal dimensions (thickness of plate, diameter of wire). Therefore, as temperature tends to zero, heat conductivity also tends to zero proportionally to the third power of temperature. By studying heat conduction at various temperatures, one detects a peculiar maximum of function $\varkappa\left(T\right)$, which indicates a change of mechanism restricting the mean free path: Umklapp processes are replaced by scattering on the crystal boundaries (Fig. 12).

The above description of heat conduction in crystals was based on the assumption that no
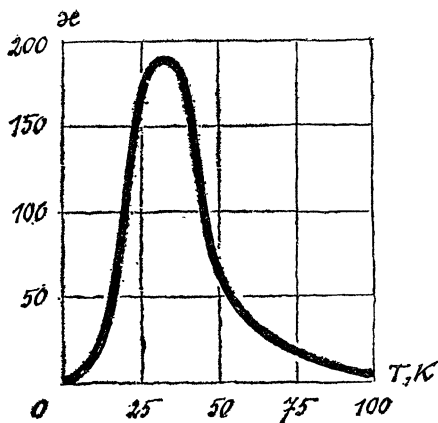
Fig. 12

other quasiparticles are present in crystals. This is true when we deal with ordinary dielectrics. In metals heat is transferred by electrons, and in magnetics, by particular quasiparticles, namely magnons. Phonons collide both with electrons and magnons, which in turn collide with phonons. Heat transfer becomes more complicated, which results in heat conductivity being a complex function of temperature. Some of these mechanisms of heat transfer will be discussed in the chapters to follow.

Is an individual phonon "observable"?

The number of phonons in a crystal is unimaginably great, even at extremely low temperatures. At a temperature one tenth of the Debye temperature, approximately 10^{20} phonons are present per cubic centimetre of a solid. Phonons

determine most of the thermal properties of solids; by studying solids physicists thereby investigate the properties of phonons. Nevertheless, it would be desirable to "touch", to "look at" a single phonon. Can this be done in principle? Of course, a phonon cannot be taken out of a crystal, because a phonon represents the motion of the atoms in the crystal. Therefore, we have to be content with studying the properties of the crystal itself.

Since low-momentum phonons are ordinary sound waves, no extraordinary techniques are required to study such phonons. In other words, by studying the sound propagation in crystals, we at the same time study the properties of individual phonons.

The study of the resonance absorption of light by a crystal can be said to represent the oldest technique of studying optical phonons. In terms of phonons this phenomenon means: a photon is transformed into a phonon. This process can be shown in the simplest drawing:

$$\mathbf{p}, \varepsilon \qquad \hbar\mathbf{k}, \hbar\omega$$

$$\mathbf{p} = \hbar\mathbf{k}, \quad \varepsilon = \hbar\omega.$$

The wavy line represents a photon with momentum $\mathbf{p}$ and energy $\varepsilon = cp$, where c is the speed of light; as usual, conservation laws are given below the drawing. The speed of light being very high, the optical phonon thus created* has a very

---

* This process never creates acoustic phonons, and not all optical phonons, because some vibrations of atoms are not related to electromagnetic waves.
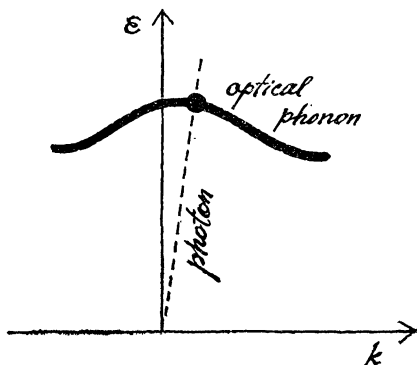
Fig. 13

low momentum (Fig. 13). Optical studies "visualize" only selected phonons.

The inelastic scattering of neutrons in crystals provides very important information on phonons. In its motion through a crystal, a neutron makes the atoms "swing", that is, excites sound waves in the crystal and thereby somewhat lowers its own energy. In terms of phonons, a neutron is said to have created phonons. The longer the path traversed by the neutron inside the crystal the higher the probability of phonon creation. As a rule, the path is short and either the neutron has not enough time to create phonons (elastic scattering) or it creates only one phonon. But this means that by studying the inelastic scattering of neutrons *we can directly measure the dispersion law of phonons.*

Indeed, as follows from the laws of momentum and energy conservation, the change in the momentum of the scattered neutron equals the phonon's momentum, and the corresponding

144

change in the neutron's energy is equal to the phonon's energy. Measure both—and you will know the dispersion law of phonons. Of course, this proposition is simple only on paper. In reality, the study of the inelastic scattering of neutrons requires sophisticated experimental equipment, high-intensity neutron sources, sensitive recording equipment, and computerized data processing. However, "the game is worth the candle". The knowledge of the phonon's energy spectra represents a deep penetration into the "construction" of solids.

Chapter 3

Two Statistics

Statistical physics—the science about collectives of atomic particles

The laws governing the behaviour of a "collective" of particles are not easily derived from the known laws of motion of individual particles. This fits in with our own experience. The prediction of mob behaviour is not easily derived from the behaviour of an individual.

The behaviour of an ensemble cannot be described in the same terms as that of an individual particle. For instance, we cannot trace the path of each particle. The number of particles is far too great for this. One has to resort to statistical methods: to calculate the mean velocity, mean energy, etc. For more than a century now this has been the subject of a special branch of physics: *statistical physics*, or simply *statistics*. But relatively recently, about thirty years ago, this field was invaded by the ideas of quantum mechanics, and a new science was born, *quantum statistics*, dealing with ensembles of particles. Quantum statistics did not invalidate classical statistics; it has delineated the proper place occupied by the latter. We shall discuss it later, but here, anticipating a little, we shall remark that as a rule classical statistics, being the limit-

ing case of quantum statistics, is valid at high temperatures. Low temperatures are the unchallenged realm of quantum statistics.* In fact, the word "low" needs serious qualification. If, for instance, we are interested in the properties of the matter in stars, then a temperature of hundreds of thousands degrees would have to be considered "low" ... .

The simplest object for physical statistics to handle is a gas, that is, an ensemble of weakly interacting, infrequently colliding particles. However, if we decide to deduce the laws of quantum statistics by studying the behaviour of a real gas, success is by no means guaranteed; as a rule, the gas will turn to liquid and then to solid before quantum effects become significant.

A gas of quasiparticles is a spectacular exception. Such gases exist in a wide range of temperatures and as a rule their quantum properties manifest themselves under ordinary (everyday) conditions.

Momentum space.
Distribution function

Molecules of gases move incessantly. The absolute temperature, that is, the temperature measured off from absolute zero, is known to be a

---

* Let us clarify: classical statistics is less exact than quantum statistics. However, the error is as a rule negligibly small at sufficiently high temperatures. As temperature is lowered, the error arising from the laws of classical statistics increases, and one has to switch to quantum statistics. A natural question is then in order: Why bother with inexact classical statistics when the more exact quantum statistics is available? The reason is that classical statistics is considerably simpler and the error is very often negligibly small.

147
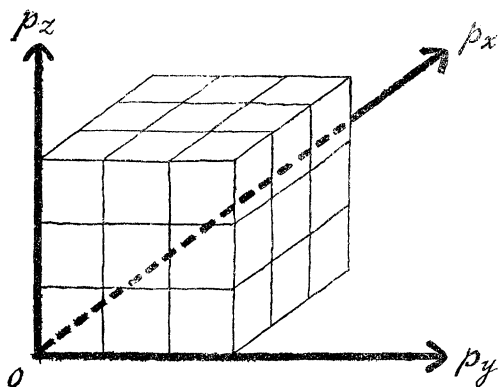
Fig. 14

measure of this molecular motion. The higher the temperature the higher the mean velocity of a molecule, while the mean kinetic energy of a molecule in the gas is simply proportional to temperature. We imagine now that we want to verify this well-known fact. We need to know how all the molecules in the gas behave. We would arrive at the required conclusion if we measured the velocity of individual molecules an enormous number of times and then calculated the mean velocity. But what graphic method can we use to display the results of measuring molecular velocities? The following method seems to be convenient.

Let us introduce the coordinate system shown in Fig. 14. The three axes are those of the momentum projections p. Any point in this three-dimensional space represents a specific value of momentum not only by magnitude but also by

direction. In particular, the origin of the coordinates represents zero momentum, that is, the momentum of a fixed particle. Now let us divide the momentum space* thus obtained into small parallelepipeds with the length of the parallelepiped edges corresponding to, say, the accuracy of momentum measurements. Now, having measured momentum, we mark a dot in a corresponding three-dimensional cell. The more often we encounter a particle with a given momentum the greater the number of dots within the corresponding cell of momentum space.

With a sufficient number of measurements and instruments with maximum accuracy, we should arrive at Fig. 15; we would find:

(a) the distribution of dots in momentum space possesses spherical symmetry;

(b) the gas of dots has maximum density close to the origin of the coordinates, that is, at a distance from the origin smaller than $\sqrt{2Mk_BT}$ (M is the mass of a particle);

(c) as we move from the origin the gas density diminishes very rapidly.

The number of dots in a cell of momentum space gives the number of particles with the prescribed value of momentum, provided the cells are made sufficiently small. Obviously, this number is a function of momentum and is called the *distribution function* since it characterizes the distribution of gas particles over momentum. The distribution function shows what portion of the

---

* This term has already been introduced. We hope that it will cause no misunderstanding. Note that momentum space is something just as objective as the ordinary space we live in, although distances in it are measured in g·cm/s.
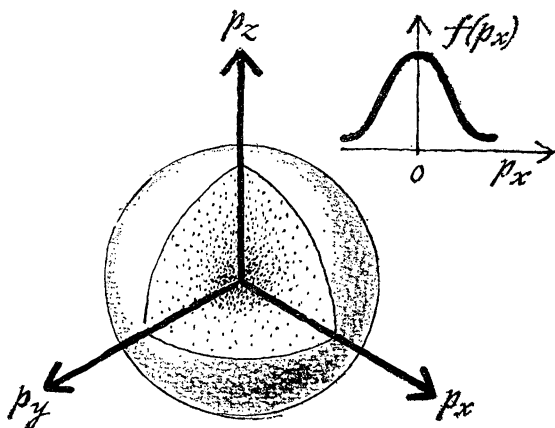
Fig. 15

total number of gas particles has a particular momentum. Hence, Fig. 15 is a particular representation of the distribution function.

By conducting our "experiment" (by measuring momenta and marking dots in momentum space) at different temperatures, we shall find that the distribution of particles is a function of temperature as well. As temperature increases, the "cloud" of dots spreads out and its density diminishes; as temperature falls, the cloud of dots contracts toward the origin and its density increases. In other words, an increase in temperature increases the number of fast particles, and a decrease in temperature increases the number of slow particles.

It might seem that Fig. 15, in which the cloud is densest close to the origin of the coordinates,
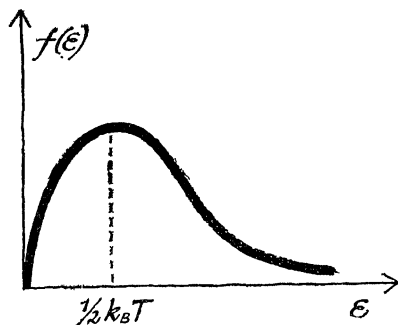
Fig. 16

contradicts our correct idea that most of the particles have energies close to $k_B T$ (their momentum is of the order of $\sqrt{2M k_B T}$). We remove the contradiction if we recall that our drawing shows the distribution of particles not only with respect to the magnitude of momentum but also with respect to direction.

By calculating the number of particles with identical energies but moving in different directions and plotting the corresponding curve, we shall find that the distribution function, being a function of energy ε, has a maximum at $\varepsilon = (1/2) \, k_B T$, according to our concepts (Fig. 16).

The distribution function is one of the most important characteristics of an ensemble of particles. Knowing this function, that is, knowing the dependence of the distribution function on momentum, we can calculate a wide variety of properties of a gas, such as its pressure, heat capacity, mean particle velocity, and in particu-

lar derive (not postulate!) the famous equation of state of an ideal gas ($PV = RT$).

All ideal gases of elementary particles are identical in classical statistics. The difference between gases observed in nature is explained only by the structure of their molecules.

We are not going to deal with this feature, because we are interested in gases of quasiparticles, elementary by definition.

Helium isotopes help understand quantum statistics

For further progress, let us describe the results of the following thought experiment (Gedanken experiment, in Einstein's terminology). We fill two identical containers with the same number N of helium atoms. The container volume is V. However, one container is filled with He^4 atoms, and the other, with He^3 atoms.

The symbol He^4 denotes the usual heavy isotope of helium, and the symbol He^3, its lighter isotope. Chemically both helium isotopes are absolutely identical, as is normal for isotopes, and differ only in that the He^3 nucleus lacks one neutron as compared with the He^4 nucleus, which makes it 25% less massive. There is another important difference between He^4 and He^3 atoms, but this will be discussed later.

Let us plot the distribution function of helium atoms at different temperatures. From the standpoint of classical statistics, we must expect identical results from measurements in each container. However, as the temperature falls we notice a gradually increasing difference between the distribution functions of He^4 and He^3 atoms.

152

Actually, this difference in the properties of helium isotopes becomes significant only at a very, very low temperature—at a temperature of several hundredths of one degree (on the Kelvin scale). This figure shows the ideal character of the Gedanken experiment described above. Long before we reach a temperature of the order of 10^{-2} K, helium will have become liquid, and the experiment will have to be stopped. However, this is a thought experiment!

We have just mentioned the temperature at which the difference in the properties of He^4 and He^3 becomes appreciable. How did we determine it? We shall give a formula whose origin will be indicated a little later.

Quantum effects (and obviously, all differences in the behaviour of He^4 and He^3 are caused by quantum effects) become appreciable when a gas is cooled to temperatures of the order of or below the value

$$T_0 = \frac{1}{k_B} \frac{\hbar^2}{2M} \left(\frac{N}{V} \right)^{2/3} .$$

Temperature T_0 has been given a rather strange name. It is called the *degeneracy temperature*. At temperatures below the degeneracy temperature the gas is said to be *degenerate*. The estimated temperature given above ($T_0 \approx 10^{-2}$ K) is obtained if we substitute 2.7×10^{19} cm^{-3} for N/V, and the mass of the helium atom for M. The chosen value of N/V corresponds to atmospheric pressure at room temperature. In other words, we began the experiment at room temperature, filling the containers with helium at atmospheric pressure.

But where did we get the above formula from?

In order to understand its origin we shall have to turn to the uncertainty relation (see formula (11)).

The mean distance between particles in a gas is $(N/V)^{-1/3}$; this gives the uncertainty in coordinate, $\Delta x \approx (N/V)^{-1/3}$. Therefore, the uncertainty in momentum is $\Delta p_x \gtrsim \hbar \, (N/V)^{1/3}$. The mean energy of a particle in a classical gas, that is, in a gas at high temperature, is equal to $(3/2)k_\mathrm{B}T$; therefore, the mean momentum is of the order of $\sqrt{Mk_\mathrm{B}T}$. As the temperature decreases, the mean momentum diminishes and the minimum uncertainty in momentum $(\Delta p_x)_{\min} \approx \hbar \, (N/V)^{1/3}$ becomes independent of temperature. There is such a temperature at which the uncertainty in momentum becomes greater than the mean momentum. At a certain temperature the mean momentum *is equal to* $(\Delta p_x)_{\min}$; this is precisely the degeneracy temperature, T_0. It is clear that quantum effects cannot be ignored if the temperature is close to T_0 and all the more so if it is below T_0.

Bose-Einstein condensation

Let us continue our examination of helium isotopes. As the temperature falls, the distribution function of He^3 atoms over momentum will gradually change, differing more and more from the classical function. A little later we shall discuss the form this function takes at absolute zero and at temperatures well below T_0.

Let us turn now to the heavy helium isotope. Experiments conducted at low temperatures (and now we shall call low those temperatures that are below the degeneracy temperature)

demonstrate to us a striking phenomenon that will make us repeat the experiments dozens of times. Helium molecules disappear from the closed container! They disappear *without leaving* the container! Let us describe the experiment in a little more detail before we try to find an explanation of this effect, an effect which could undermine our confidence in the impossibility of matter disappearing.

Beginning at temperature $T_{\mathrm{BE}} \approx 4.4 T_0$, the number of gas molecules in the container (we determine this number by counting all the moving molecules) begins to dwindle, and one even can find the temperature dependence of this number: $N_T = N (T/T_{\mathrm{BE}})^{3/2}$. At absolute zero no particle can be found ... . What is going on? An attentive reader may have noticed the message in the words put in parentheses: "we determine this number by counting all the moving molecules". Indeed, here lies the answer to the puzzle. We observe only moving molecules. At high temperatures no particles are at rest (to be more precise, the number of such particles is *infinitely small*), but beginning at temperatures T_{BE}, a *finite number* of particles has zero momentum*. At absolute zero all particles of gas are in this state. One should not think that particles with zero momentum *lie* on the bottom of the container. We recall the uncertainty relation. Once the momentum of these particles is definite (equal to zero), their coordinate (i.e. their position

---

* According to classical statistics, the number of particles with zero momentum is' infinitely small at any temperature, however low it might be, and only at $T = 0$ are all particles at rest.
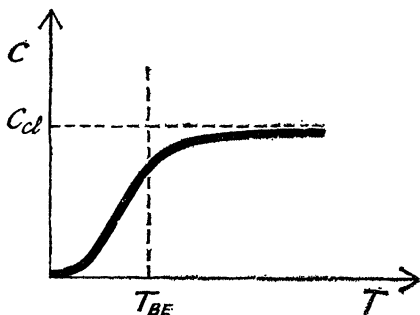
Fig. 17

within the container) becomes completely uncertain. The molecules are somewhere within the container, but where? Not only is this question unanswerable, but it just cannot be posed— it is meaningless.

The phenomenon described in the above paragraph is called the *Bose-Einstein condensation*. It has very little in common with usual condensation, such as the condensation of a gas into a liquid.

One could suspect that the Gedanken experiment was specially designed to demonstrate the apparent disappearance of particles. Obviously, we could weigh the container each time and ascertain that all the particles were still within the container. However, all other methods of studying the gas, with the exception of weighing it, would definitely demonstrate the disappearance of particles. In this sense, the best demonstration is provided by measuring heat capacity. A sharp drop in heat capacity C below T_{BE}
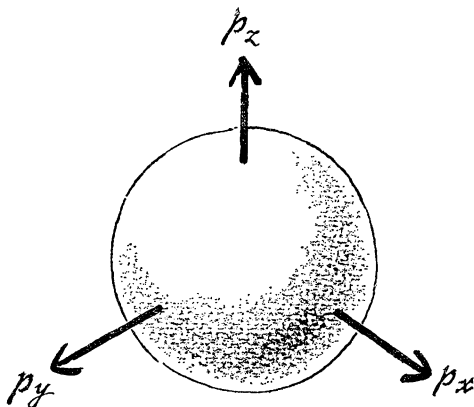
Fig 18

is a corollary of the Bose-Einstein condensation of particles, that is, of their transition to the state of zero momentum (Fig. 17).

Fermi-Dirac degeneracy

At low temperatures the distribution function of He³ atoms over momentum differs so drastically from the distribution function of a heavy-isotope atom that one simply tends to wonder whether these atoms represent the same chemical element. The difference is especially dramatic at absolute zero. Indeed, at $T = 0$ He³ atoms fill a sphere in momentum space (Fig. 18). In memory of Enrico Fermi, it is referred to as the *Fermi sphere*. The higher the density of particles the greater the radius of the Fermi sphere. The energy of particles located on the surface of the Fer-
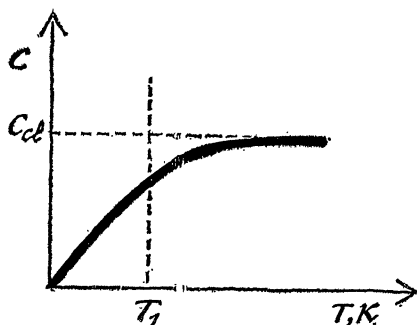
Fig. 19

mi sphere, ε_F (also referred to as the *Fermi ener-gy*), approximately equals the degeneracy tem-perature multiplied by the Boltzmann constant, and the mean energy of a particle of gas is equal to $(3/5)\varepsilon_F$, which is a large quantity on the atom-ic scale. The temperature $T_F = \varepsilon_F/k_B$ is often called the *Fermi temperature*.

Of course, this peculiar distribution of par-ticles over momentum results in unique proper-ties of the gas as a whole. If we measure (a thought measurement, since liquefaction is, as before, ignored) the heat capacity of He^3 for a wide range of temperatures, the results will be the curve shown in Fig. 19. Note the characteristic linear dependence of heat capacity on temperature as we approach absolute zero. Figure 20 plots the *distribution* of He^3 atoms *over energies* at absolute zero temperature and at a nonzero temperature low when compared with the Fermi temperature $(k_B T \ll \varepsilon_F)$.

158
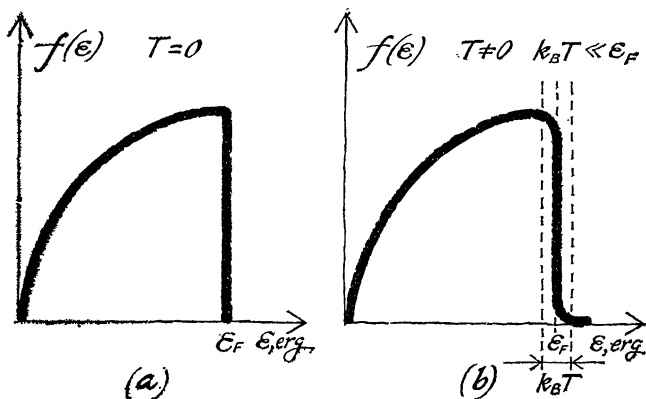
Fig. 20

We shall see later that the behaviour described above of the lighter helium isotope is a general property of some gases and is referred to as the *Fermi-Dirac degeneracy*. We also want to emphasize that everything said in this section directly concerns the electron gas.

Fermions and bosons

The difference between the two helium isotopes is so drastic that a natural question about its origin immediately arises. What is the reason for it? Is it simply a matter of difference in masses? No, this is not so. There exist only *two* manners of behaviour, or, if we resort to strict language, two statistics. In memory of the scientists who discovered them, these are called the *Bose-Einstein statistics* (among the particles that obey this statistics are He^4 atoms) and the *Fermi-Dirac sta-*

tistics (He³ atoms, for example). A gas of particles obeying the Bose-Einstein statistics is called a *Bose gas*, and its particles are called *bosons*. A gas of particles obeying the Fermi-Dirac statistics is called a *Fermi gas*, and its particles are called *fermions.**

Above we attempted to illustrate how different the patterns of behaviour of the Bose and Fermi gases are. Do not forget, however, that this difference manifests itself only at low temperatures. At high temperatures, that is, at a temperature above the degeneracy temperature, all gases of elementary particles behave in the same manner.

What is then a microscopic property, that is, a property characterizing an individual particle, which determines the behaviour of an ensemble of particles? We have already mentioned that this property is the *spin of a particle*, that is, the intrinsic angular momentum of the particle.

The following strict rule can be formulated:

> All particles with half-integral spin obey the Fermi-Dirac statistics, and those with zero or integral spin obey the Bose-Einstein statistics.

In other words, particles with integral or zero spin are bosons, and those with half-integral spin are fermions.

The difference between the behaviour of fermions and that of bosons is explained by the fact that the former are subject to the Pauli exclusion principle (see p. 44), while the latter are not.

---

* Recall the table on p. 58 and our comment about the last column.

He^4 He^3

p — proton
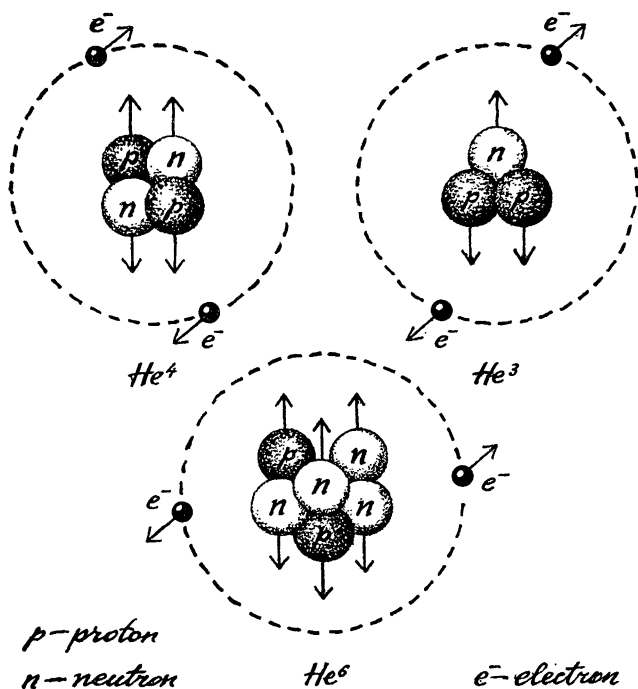n — neutron He^6 e — electron

Fig. 21

This "ban" that prohibits fermions from gather-
ing in one state is seen especially clearly in their
distribution function at temperatures close to
absolute zero. At $T = 0$ this ban prevents a
transition to the zero-momentum state for all
particles, the state so "willingly" occupied by
bosons.

The difference in the behaviour of He^4 and
He^3 atoms will be clear if we look at the sche-

matic structure of the helium atoms (Fig. 21).
The spin of He^4 atoms is zero, while that of He^3
atoms 1/2. The difference in spin appears because
He^3 lacks one neutron. The behaviour of gas
consisting of He^6 atoms is easily predictable
(such an isotope exists, though it is unstable).
Since the spin of He^6 atoms is zero, a system of
these atoms obeys the Bose-Einstein statistics.

Several formulas

This chapter is so essential for an understanding
of the book as a whole that, at the risk of antag-
onizing the reader who dislikes mathematics,
we decided to set out several formulas that in
fact express all the physics discussed in the pre-
ceding sections. Formulas require higher preci-
sion than worded statements; hence, we have to
define more accurately some of the concepts
that have already been mentioned.

Quantum statistics operates not in momentum
space but a more abstract six-dimensional *phase
space*. This space has six axes of coordinates:
x, y, z, p_x, p_y, p_z. The uncertainty relation deter-
mines unambiguously the volume of a cell in
this space. It is equal to $(2\pi\hbar)^3$.

*The mean number of particles per cell of phase
space is called the distribution function.* We shall
denote it by letter n, with a subscript BE for
bosons and FD for fermions. Although the prop-
erties of the distribution functions for bosons
and fermions are, as we have mentioned, very
different, the corresponding formulas look quite
similar. The equilibrium distribution function
for bosons, the *Bose-Einstein function*, is

$$n_{BE}(\mathbf{p}) = \frac{1}{\exp\{[\varepsilon(\mathbf{p}) - \zeta]/k_BT\} - 1}. \qquad (41)$$

The equilibrium distribution function for fermions, the *Fermi-Dirac function*, is

$$n_{\mathrm{FD}}(\mathbf{p}) = \frac{1}{\exp\{[\varepsilon(\mathbf{p})-\zeta]/k_{\mathrm{B}}T\}+1}. \qquad (42)$$

These two formulas differ only in the sign in front of unity in the denominator. In addition to already familiar quantities (energy $\varepsilon(\mathbf{p}) = p^2/2M$ and temperature T), eqs. (41) and (42) include *chemical potential* ζ. Its physical meaning will become clear a little later. Let us indicate, first of all, how to calculate the macroscopic quantities of interest such as the number of particles N in a container with volume V or their energy $\mathscr{E}$, when the distribution function is known:

$$N = \frac{gV}{(2\pi\hbar)^3} \int\int\int \frac{dp_x\,dp_y\,dp_z}{\exp\{[\varepsilon(\mathbf{p})-\zeta]/k_{\mathrm{B}}T\}\pm 1},$$
$$(43)$$

$$\mathscr{E} = \frac{gV}{(2\pi\hbar)^3} \int\int\int \frac{\varepsilon(\mathbf{p})\,dp_x\,dp_y\,dp_z}{\exp\{[\varepsilon(\mathbf{p})-\zeta]/k_{\mathrm{B}}T\}\pm 1}.$$
$$(44)$$

We integrate over the whole momentum space, $g = 2s + 1$, where s is the spin of a particle (we recall that bosons have integral or zero spins, while fermions have half-integral spins; for electrons, $s = 1/2$ and $g = 2$). A careful look at eqs. (43) and (44) reveals how they were "constructed": to calculate macroscopic quantities we have to sum the contributions to a given quantity of each particle, the summation being carried out not over particles but the cells of phase space. Factors g and V appear because the energy of a

particle is independent of the orientation of its spin and coordinate.*

The number of particles in a container is fixed (recall the conditions of the thought experiment described in this chapter). Temperature can also be considered fixed. It can be changed, though, by heating or cooling the container's walls. Hence, eq. (43) can be considered as an equation that yields the value of chemical potential ζ. The same statement defines the physical meaning of chemical potential. This potential realizes the *normalization of the distribution function.*** Chemical potential would be meaningless in a gas so "fantastic" that the number of particles in it were not conserved in a closed volume when the temperature changed.

Let us recast eqs. (43) and (44) in a simpler form, by rewriting the integrals in the spherical system of coordinates and then integrating over angle:

$$N = \int_0^\infty \frac{\nu(\varepsilon)\, d\varepsilon}{\exp\left[(\varepsilon-\zeta)/k_BT\right] \pm 1}\,,$$

$$\mathscr{E} = \int_0^\infty \frac{\varepsilon\nu(\varepsilon)\, d\varepsilon}{\exp\left[(\varepsilon-\zeta)/k_BT\right] \pm 1}\,. \tag{45}$$

The quantity

$$\nu(\varepsilon) = \frac{gVM^{3/2}}{\pi^2\hbar^3\sqrt{2}}\,\varepsilon^{1/2} \tag{46}$$

---

* We emphasize that a region in the **p**-space with volume Δ contains $gV\Delta/(2\pi\hbar)^3$ fermions (the importance of this fact will become clear in the next chapter).

** This is not a complete definition of the physical meaning of chemical potential. Unfortunately, its other roles cannot be discussed within the scope of this book.

is the *density of states per interval of energy* (ε, $\varepsilon + d\varepsilon$), and

$$n(\varepsilon) = \frac{\nu(\varepsilon)}{\exp[(\varepsilon - \zeta)/k_B T] \pm 1} \qquad (47)$$

is the *distribution function of particles over energy*. We have mentioned on numerous occasions that the laws of classical statistics are valid at high temperatures. Then how can we realize the limiting transition to high temperatures in the above formulas? At first glance, by the most improbable method: the unity in the denominator must be ignored. The point is that according to eq. (43), we obtain for $T \gg T_0$ the expression

$$\zeta \approx \zeta_{cl} = -\frac{3}{2} T \ln \frac{T}{T_0},$$

where T_0 is the very degeneracy temperature whose formula has been discussed earlier; this means that, first, $\zeta < 0$, and, second, $|\zeta| \gg T$; this enables us to ignore the unity in the denominator.

Finally,

$$n_{cl}(\mathbf{p}) = \exp\{[\zeta - \varepsilon(\mathbf{p})]/k_B T\}. \qquad (48)$$

At high temperatures the mean number of particles per cell of phase space is small, they do not "interfere" with each other, their quantum interaction (the Pauli exclusion principle, etc.) does not manifest itself, so that classical statistics is valid.

Now a few words about the Bose-Einstein condensation and the Fermi-Dirac degeneracy. As shown by the formula, the chemical potential of a Bose gas cannot be positive. The number of particles n_{BE} in a cell of phase space for which $\varepsilon(\mathbf{p}) < \zeta \neq 0$ would be negative when $\zeta > 0$,
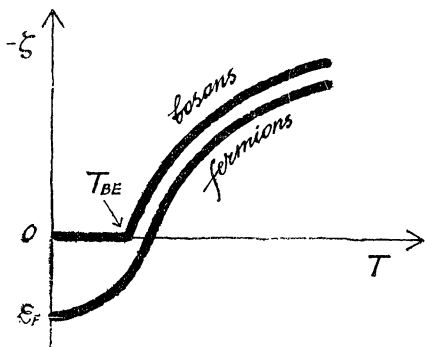
Fig. 22

and would be infinite, $n_{\text{BE}} = \infty$, when $\varepsilon\,(\mathbf{p}) = \zeta \neq 0$! However, according to eq. (43) or (45), ζ can be zero. It is not difficult to find the temperature T_{BE} at which this takes place if we assume $\zeta = 0$ in the expression for N:

$$N = \int\limits_{0}^{\infty} \frac{\nu\,(\varepsilon)\,d\varepsilon}{\exp\,(\varepsilon/k_{\text{B}}T_{\text{BE}}) - 1} . \qquad (49)$$

An estimate for T_{BE} has been given already (see p. 155). But what will happen at $T < T_{\text{BE}}$? The number of *moving* particles, as we know, is not conserved in this case. As the temperature falls, their number decreases, and hence $\zeta \equiv 0$ (Fig. 22). So we find that the fantastic gas does exist, after all. I do not think that an understanding of this phenomenon makes it any less fantastic.

Understanding the Bose-Einstein condensation helps to explain the nature of superfluidity and superconductivity.

The chemical potential of a Fermi gas can have any sign: at temperatures close to T_0, it reverses its sign. But what will happen at $T \to 0$? Let us carefully calculate the limit

$$\lim_{T \to 0} \frac{1}{\exp\left[(\varepsilon - \zeta)/k_{\mathrm{B}}T\right] + 1}$$

assuming that at $T = 0$ the quantity ζ is neither zero nor infinitely large (let $\zeta_{T \to 0} = \zeta_0$). Then

$$\lim_{T \to 0} \frac{1}{\exp\left[(\varepsilon - \zeta)/k_{\mathrm{B}}T\right] + 1} =$$
$$= \begin{cases} 1 & \text{for} \quad \varepsilon < \zeta_0, \\ 0 & \text{for} \quad \varepsilon > \zeta_0. \end{cases} \tag{50}$$

This discontinuous function is called the *Fermi step function*. It shows that at absolute zero chemical potential coincides with the Fermi energy ($\zeta_0 = \varepsilon_{\mathrm{F}}$). Figure 20 shows the Fermi-Dirac function multiplied by the density of states at $T = 0$ and at $T \ll T_{\mathrm{F}}$. Now it is easy to find, by using the first of eqs. (45) (of course, the one with the plus sign) and also eq. (46), that

$$\varepsilon_{\mathrm{F}} = \left(\frac{6\pi^2 N}{gV}\right)^{2/3} \frac{\hbar^2}{2M}. \tag{51}$$

Let us return to photons and phonons and compare eqs. (19) and (34) with the formulas in the present section. We find that *photons and phonons are bosons*. Naturally, their chemical potential is zero, since the number of photons is the greater the higher the temperature of the cavity walls (this cavity emits photons), and the number of phonons is the greater the higher the crystal temperature. The numbers of photons and phonons are not conserved.

Chapter 4

Electrons

Electron energy spectrum of crystals

The influence of quantum mechanics is already felt in the combination of words "energy spectrum". The word "spectrum" comes from the wave side of things and the word "energy" comes from the corpuscular side of things. The discussion to follow of the electron properties of crystals has several goals:

(a) to understand why some crystals, the so-called conductors, conduct electric current well, while other crystals, insulators, conduct electric current poorly;

(b) to explain the existence of materials with intermediate properties: semiconductors, semimetals;

(c) to analyze, using a number of examples, how quantum mechanics and quantum statistics explain the behaviour observed in experiment.

Our exposition of the electron properties of solids will not follow the historical evolution of the relevant concepts (just as we did not follow it in the preceding chapters). We shall try to describe the electron properties of solids in terms

of today's concepts. As mathematical techniques are taboo in this book, we shall resort to analogy, intuition's assistant. We must confess that this chapter will require more effort on the part of the reader than the preceding chapters.

An electron in a crystal is in a force field created by ions which are arranged in space periodically. This phrase is already a "quantum" phrase. A classical electron is found *either* in the field of one ion *or* in the field of another ion. A quantum electron, however, does not fear potential barriers and can move from ion to ion with relative freedom. As far as this property is concerned (in scientific language, delocalization) an electron in a crystal resembles a free electron. The energy of a free electron, ε, is a function of momentum* $\mathbf{p}$. In order to avoid the mathematical techniques of quantum mechanics, we will use the de Broglie relation in its initial form: $\hbar\mathbf{k} = \mathbf{p}$, where $\mathbf{k}$ is, as before, the wave vector. When we were speaking about the structure of crystals at the beginning of the preceding chapter, we emphasized that a periodic crystal "creates" reciprocal space ($\mathbf{k}$-space) that is also periodic. Then, in the chapter about phonons (p. 133-4), we transformed it into the periodic $\mathbf{p}$-space. All quantities dependent on momentum are periodic functions of $\mathbf{p}$. Therefore, the energy of an electron in a crystal, ε, is a periodic function of $\mathbf{p}$.

Let us now tackle another side of this problem. An electron in an isolated atom is in a quite definite state. Let us represent the energy of this state by a horizontal line on which the electron

---

* In the case of a free electron, $\varepsilon = p^2/2m = mv^2/2$, where m is the electron's mass; $\mathbf{p} = m\mathbf{v}$, where $\mathbf{v}$ is the electron's velocity.
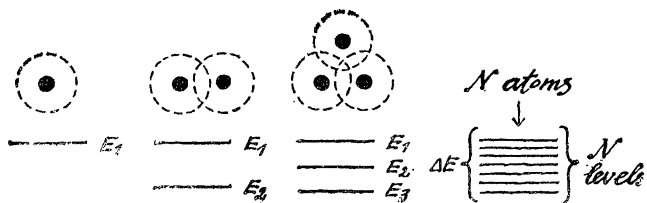
Fig. 23

can "perch" (Fig. 23). Now we can bring the two atoms, each with its own horizontal line, closer. What will happen to these horizontal lines when the atoms are close together? A quantum mechanical calculation shows that the lines will move slightly apart. In the figure they are placed one *on top of* the other since the electron is "present" in both atoms in each of the two states simultaneously (each line denotes the energy of the state).

It is time now to say "and so on". When N atoms are brought together, each of the states is split into N states. And there is no necessity for all of the atoms to be close to one another: an electron will reach any one of them by virtue of tunnelling. The maximum splitting will be determined by the closest neighbours. Tremendous number of atoms is united in a crystal: about 10^{23} per 1 cm^3. As a result, each atomic state in a crystal is split into a band of allowed values of energy. Let us enumerate these bands: 1st band, 2nd band, ... , jth band, ... Consequently, when describing the state of an electron in a crystal, we have to indicate the *number of the band* to which it belongs; but this is obviously insufficient. Can the position of an electron within the

band be given with more precision? Bands occur owing to the ability of the electron to move from atom to atom, so it is not surprising that the state of the electron in the band is a function of momentum **p**; as we have said, this is a periodic function. In order to describe the state of the electron more precisely, we also have to indicate the direction of its spin: "up" ($s_z = 1/2$) or "down" ($s_z = -1/2$). The following notation is often used to represent the state of an electron in a crystal: $|j, \mathbf{p}, s_z\rangle$; it should now be clear to the reader.

An electron in a crystal resembles both an electron in an atom and an electron in a force-free space (equipotential space). What is essential here is the resemblance with a free electron. Therefore we set out the necessary information in a table. It does, of course, require some explanation, but since this table is a result of an exact theory, here we shall accept it as given.

	Free electron	Electron in crystal
State	$\|\mathbf{p}, s_z\rangle$	$\|j, \mathbf{p}, s_z\rangle$
Energy	$p^2/2m$	$\varepsilon_j(\mathbf{p}) = \varepsilon_j(\mathbf{p} + 2\pi\hbar\mathbf{b})$
Velocity	$\mathbf{v} = \mathbf{p}/m$	$\mathbf{v}_j = \partial\varepsilon_j(\mathbf{p})/\partial\mathbf{p}$
Equation of motion	$d\mathbf{p}/dt = \mathbf{F}$	$d\mathbf{p}/dt = \mathbf{F}_{\text{ext}}$ *
Law of momentum conservation in collision of two electrons	$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_1' + \mathbf{p}_2'$	$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_1' + \mathbf{p}_2' + 2\pi\hbar\mathbf{b}$

* The subscript "ext" indicates that the electron's momentum in the crystal is determined by an external force and not by the total force. After creating the state of the electron, $|j, \mathbf{p}, s_z\rangle$, the forces due to the ions are, so to say, switched off. This is an important fact: $\varepsilon_j(\mathbf{p})$ can often be considered the *kinetic energy* of electrons in a crystal. But in fact it is the *total energy* in the absence of external forces.

To summarize, an electron in a state $|j, \mathbf{p}, s_z\rangle$ possesses energy $\varepsilon_{j, s_z}(\mathbf{p})$. The arguments j, s_z, and $\mathbf{p}$ are written in a different form in order to emphasize that j and s_z are discrete, while vector $\mathbf{p}$ assumes an arbitrary value. If we limit the analysis to physically nonequivalent states, it is sufficient to use $\mathbf{p}$ from the first Brillouin zone (see Ch. 2, p. 113). Functions $|j, \mathbf{p}, s_z\rangle$ and energies $\varepsilon_{j, s_z}(\mathbf{p})$ can be calculated by using particular assumptions concerning the interaction between the electrons and ions. This will be discussed later, although only in passing. Even if $\varepsilon_{j, s_z}(\mathbf{p})$ are known functions, still we know very little about the properties of the crystal. As little as about a company in which one has read the list of staff positions but does not know the people who occupy these positions. We need to know how the electrons occupy the states $|j, \mathbf{p}, s_z\rangle$.

We consider a crystal with volume V composed of $\mathfrak{N}$ cells, that is, $\mathfrak{N} = V/a^3$. Let each cell of the crystal contain Z electrons (in the case of elementary crystals, Z is the atomic number of the component atoms). All in all, the crystal contains $\mathfrak{N}Z$ electrons. Let us distribute these electrons over the bands, assuming that the energy of an electron is independent of the projection of its spin ($\varepsilon_{j, 1/2}(\mathbf{p}) \equiv \varepsilon_{j, -1/2}(\mathbf{p})$). Then *two* electrons can be placed in each state. Let us fill the band by using a method that has already been used to fill atomic energy states: to minimize the energy and to satisfy the Pauli exclusion principle. Since the volume of the first Brillouin zone is equal to $(2\pi\hbar)^3/a^3$, then *all* states are filled
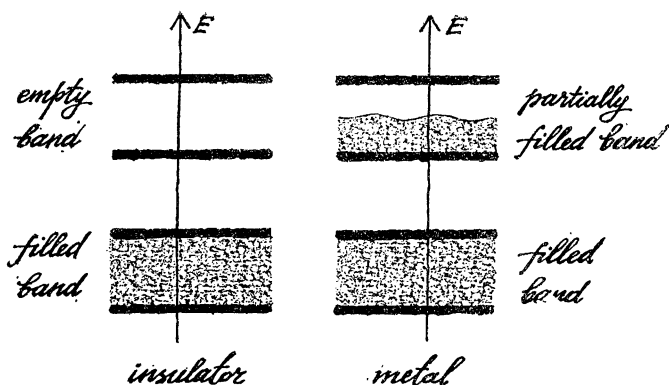
Fig. 24

for $2V/a^3 = 2\mathfrak{N}$ electrons in the zone (see p. 113
and the first footnote to p. 164). Consequently, an
integer number of bands can be filled if Z is even,
while one band is invariably half-empty (or,
shall we say, half-filled?) if Z is odd.

The answer to the question chosen as the head-
ing for this section depends crucially on wheth-
er the material possesses an empty band or a
partially filled band.

Let us assume that an electric field is applied
to the crystal. An electron accelerated by this
field increases its energy, that is, passes to a high-
er energy level. This process is only possible,
however, *if either there is such a level or if it is
unoccupied*. This condition is satisfied in crystals
with only partially filled bands, and is not satis-
fied in crystals in which the bands are either com-
pletely occupied or empty: allowed unoccupied
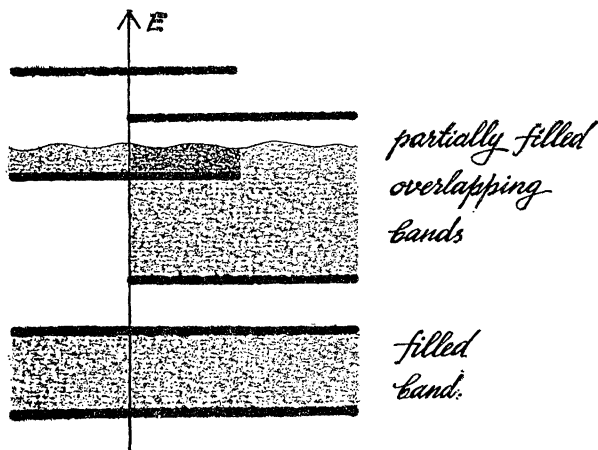energy levels are separated by a gap from occu-

Fig. 25

pied levels (Fig. 24). It might seem that a simple rule can be formulated: even Z means an insulator (dielectric) while odd Z means a conductor (metal).

Let us confront this hypothesis with facts by having a look at the first squares of the Mendeleev table. Hydrogen is a dielectric consisting of H_2 molecules, two electrons per cell, one filled band; helium is similar: a monatomic crystal (under pressure), two electrons per cell, a dielectric; lithium has three electrons per cell, a metal. So far, the experiment bears out the theory. However, the "beautiful" picture is shattered by beryllium: it has four electrons per cell, but is a metal and a good conductor. The problem is that we have oversimplified the scheme: the allowed energy bands can overlap (Fig. 25).

174

Therefore there is no simple criterion based on evenness or oddness of Z. All we can say is that if a crystal possesses bands partially occupied by electrons, it is a conductor, a metal; if it has no such bands, it is an insulator, a dielectric. The overlapping of bands is fairly common. Conductors (metals) are encountered among elements more frequently than insulators.

Semiconductors, semimetals

The preceding section has clearly showed that metals at absolute zero can conduct electric current; we shall see later that this capability deteriorates as the temperature increases, that is, the electrical resistance increases.

In addition to metals, there exists another class of crystalline electron conductors. This class includes such elements as silicon, germanium, selenium, and a number of compounds, such as cuprous oxide, and so on. All these materials conduct electric current much worse than metals and therefore are called *semiconductors*. In addition, they may become dielectric, that is, totally nonconducting, at absolute zero. In contrast to metals, their conduction is thermal in nature. This means the following.

We assume that the bottom of the empty energy band (called the *conduction band*) lies very close to the top of the completely filled band (called the *valence band*). Then at a sufficiently moderate energy of thermal excitation the electron will jump from the completely filled lower band to the empty upper band. At absolute zero electrons are not excited thermally, no electrons lift to the upper band, and the crystal will be dielec-

175

tric. As the temperature increases, the number of electrons lifted to the upper band will increase. Naturally, the conducting property will thereby be improved, that is, the resistance will diminish.

It should not be thought, however, that only electrons transfer the electric charge in the conduction band of a semiconductor. Having left the valence band, an electron leaves behind an empty state. Hence, the remaining electrons in the lower band can now be accelerated and transferred to the upper states. Obviously, if in this transfer an electron moves to the right, the empty state (hole) moves to the left (like a bubble in water, when in fact the centre of gravity of the mass of water moves downward). We know that electrons move towards the anode, that is, to the positive electrode. Consequently, holes move towards the cathode and thus behave as positive charges.

The sign of the electric charge is not the only difference between the holes and electrons. Their respective contributions to the total electrical conduction are different (the electron's current is not simply one half of the total current) because holes and electrons transfer the electric charge in the external field with different velocities. The contribution of electrons can be experimentally distinguished from that of holes. For this it is sufficient to put the semiconductor into the magnetic field perpendicular to the electric current flowing through the semiconductor. The electrons will be deflected in one direction, and the holes, in the opposite direction. The deflection of charges in the magnetic field produces an additional potential difference in the di-

rection normal both to the current and to the field (the Hall effect*). This potential difference can be measured by a galvanometer. The direction of the galvanometer's deflection indicates whether electrons or holes predominate in the sample. The Hall voltage may be of different sign even in metals. This demonstrates that even in metals holes can take part in electrical conduction.

So far we have been speaking about semiconductors containing both holes and electrons. It is known, however, that most semiconductors possess either electron or hole conduction. Experiments show that all such semiconductors contain impurities of foreign elements. What is the role played by the impurities?

The presence of impurities in a crystal means that the crystal is no longer ideal. The theory (the band theory) shows that nonideal crystals have in their energy spectra both bands of allowed values of electron energy and individual levels located in the gaps between the bands (Fig. 26). These levels can be either occupied or vacant. Let us assume that the occupied levels are very close to the bottom level of the empty band. Thermal motion will then force some of the electrons to cross into the empty band (conduction band), thus giving a semiconductor with purely electron conductivity.

In the case under discussion, the hole energy levels correspond to the states related to the nuclei; the hole conduction can then be realized by electrons hopping in from the neighbouring atoms. Sometimes this is indeed the case, and the cor-

---

* The Hall effect will be discussed in detail later (see Chapter 5).

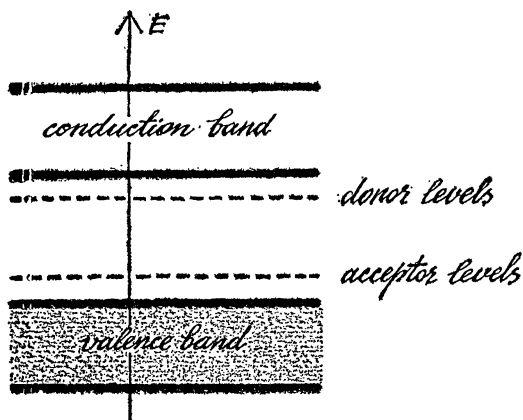conduction band

donor levels

acceptor levels

valence band

Fig. 26

responding conduction is called "hopping" con-
duction. But as a rule, either all impurities lose
their easily mobile electrons or the spacings be-
tween impurity atoms are too great, so that the
probability of hops is low.

When the empty levels are close to the top
level of the filled valence band, then thermal mo-
tion excites some of the electrons to these levels,
and holes appear in the valence band. This pro-
duces a hole semiconductor. The electrons are
then connected to the atomic nucleus.

When a semiconductor is illuminated with vis-
ible or ultraviolet light, some electrons absorb
energy from the light and can move from the
valence band to the conduction band, leaving
behind mobile holes. This means that the semi-
conductor is capable of conducting an electric
current while being illuminated, even at a very
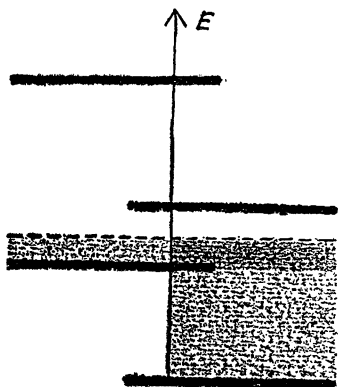low temperature. This phenomenon is called
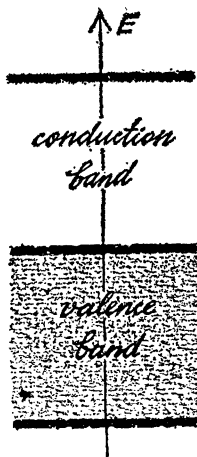*photoconductivity.*

Fig. 27   Fig. 28

If a semiconductor is illuminated with *monochromatic* light, then photoconductivity can be used to measure the spacing (in energy units) between occupied and unoccupied states. By using Bohr's formula, we easily obtain $\hbar\omega = \Delta\varepsilon$, where $\Delta\varepsilon$ is the change in the electron's energy. This is one of the fundamental methods of studying the energy structure of semiconductors.

It is customary now to single out another class of electron conductors, namely *semimetals*. Semimetals are metals, but metals which only just escaped being dielectrics or rather semiconductors. A characteristic example of a semimetal is bismuth. Bismuth has two atoms per unit cell. It could be a dielectric, but its valence band and conduction band slightly overlap (Fig. 27).

179

The overlapping is truly tiny. The density of electrons (or of holes) in bismuth is smaller than in ordinary "good" metals by a factor of 10^5.

Semiconductors and semimetals lie between metals and dielectrics, but right on the border we find *zero-gap semiconductors* in which the valence band and the conduction band are in contact, as shown in Fig. 28. Such materials have been found and are now actively studied.

Fermi surface

This chapter will hereafter deal only with metals and their properties. Consequently, we shall speak about crystals with one or several partially filled bands. Let us try to picture the momentum space of such a metal. This space is partially filled with electrons. The surface which separates the region partially occupied by electrons from the region free from electrons is called the *Fermi surface*, analogous to the Fermi sphere of the free gas of fermions. The Fermi surface is one of the constant-energy surfaces* $\varepsilon_j(\mathbf{p}) = \varepsilon$. Since $\varepsilon_j(\mathbf{p})$ is a periodic function of momentum, the Fermi surface, like other constant-energy surfaces, can have a simple shape (sodium and potassium are two examples) or a very complex shape.

Figure 29 shows some of the actual shapes of the Fermi surfaces. For the sake of clarity we plot the Fermi surface only within the first Brillouin zone and rely on the reader's imagination for the periodic extrapolation of the structure.

---

* The constant-energy surface is a surface of fixed energy. For free particles, constant-energy surfaces are spheres with radius $p = \sqrt{2M\varepsilon}$.

In one case we lend a helping hand: compare Figs. 29 and 31, which both show the Fermi surface of copper.

As follows from the fact that inside the Fermi surface, or rather within that part of the surface that lies within the first Brillouin zone, we find all the electrons from the partially filled band (the so-called conduction electrons, with density n_e per 1 cm^3). There exists a simple relation

$$n_e = \frac{2\Delta_F}{(2\pi\hbar)^3}. \qquad (52)$$

Here Δ_F is the volume of the Fermi surface within the first Brillouin zone (see the first footnote to p. 164). In some situations the Fermi surface looks as if it were turned inside out. In such cases it is more convenient and more typical to relate the volume of the Fermi surface to the number of empty states in the band. As in semiconductors, these empty states are called holes, and the Fermi surface is referred to as the hole surface. Again, as in semiconductors, holes in metals make a contribution to the Hall effect of the sign opposite to that of the conduction electrons.

The Fermi surface is a very important characteristic of electrons in metals. The details of the structure of the Fermi surface determine many subtle properties of metals, especially their behaviour in magnetic fields.

The reconstruction of the shape of Fermi surfaces in a number of metals has been achieved by measuring their resistance in magnetic field and the absorption of radio waves in the centimetre and millimetre wavelength range by the surface of metals, by probing metals with ultrasound waves, by plotting the magnetic moment of the
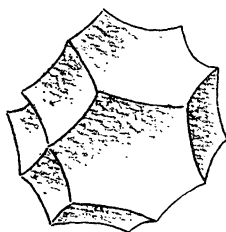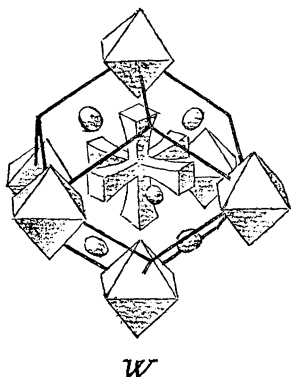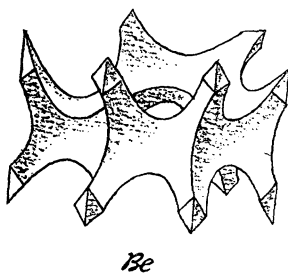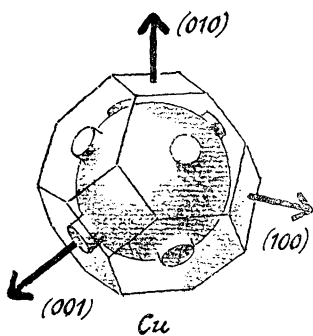
Cu

Be

W

Al, In

Fig. 29   Fig. 30

metal as a function (often a complicated func-
tion) of the magnetic field. The experiments were
conducted under extreme conditions: at tempera-
tures close to absolute zero, in very strong mag-
netic fields, and in a range of frequencies achieved
only with tremendous effort. Now there are
handbooks of Fermi surfaces, detailed study of
them continues, and the range of metals covered
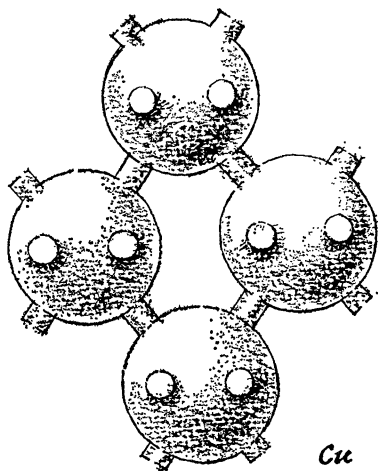is complete.

182

Cu

Fig. 31

As the electron energy spectrum was being studied, it became clear, at least in principle, why some Fermi surfaces are very simple (Na, K), while others are very complicated.

Let us forget the ions for a moment. Now a conduction electron moves in the empty space, where the motion of it is very similar to the propagation of a wave. Let us imagine waves in the ocean. The ions in the lattice are piles driven into the ocean's bottom. A wave is slightly reflected by each row of piles. The reflected waves are mutually amplified if the reflection is each time in the same phase. The amplitude of the wave reflected by the lattice of piles soon becomes equal to that of the incident wave. Obviously, this prevents the propagation of any individual

wave since a wave with the same amplitude travels in the opposite direction. In fact, we cannot distinguish any more between an incident travelling wave and reflected wave, and so obtain a single resultant wave which is said to be a standing wave. A standing wave is obtained if the reflection from the consecutive rows of piles is always in the same phase; this requires that the wavelength of the travelling waves be either equal to twice the spacing between piles or be less than this distance by a factor of two, three, four, and so on.

We find, therefore, that the wave pattern is considerably transformed as compared with the free particle pattern if the waves interfere, but that the effect of ions (piles) is not so great if there is no interference. Let us formulate, in corpuscular terms, the condition of, this time, the impossibility of interference. Interference is impossible if the electron's momentum is below $\pi \hbar / a$. Here a is the minimum spacing between the "planes of the reciprocal lattice".

Momentum is maximum for Fermi electrons, that is, it equals $p_{\mathrm{F}} = \hbar \, (3\pi^2 n_{\mathrm{c}})^{1/3}$*. The condition forbidding interference, and at the same time stipulating the simple shape of the Fermi surface in cubic crystals, is written as follows:

$$ n_{\mathrm{c}} a^3 < \frac{\pi}{3} . $$

In alkali metals (with the exception of lithium) this condition is satisfied. The respective Fermi surfaces are spheres.

Before concluding this section, it remains for

* This expression is derived from eq. (52). In the case of free electrons, $A_{\mathrm{F}} = (4/3) \, \pi p_{\mathrm{F}}^3$ (see also eq. (51)).

us to understand why the ions can be treated as small perturbations (this assumption is fundamental to the above arguments). The main reason is that their comulative action is neutralized by electrons, so that the net charge of the metal is zero.

Conduction electrons form a degenerate Fermi gas

What we said at the end of the preceding section of the comparatively weak effect produced by ions on the motion of electrons justifies our treating them as particles moving in free space.*
The main role played by ions is to prevent electrons from leaving the crystal. Additional energy has to be transferred to an electron to extract it from the metal. If this is realized by the absorption of a photon, then we observe the photoeffect, as was mentioned in Chapter 1.

Here and in the next chapter we consider electrons as ordinary particles, with energy ε being equal to $p^2/2m$ and $m \approx 10^{-27}$ g. Electrons are described by the Fermi-Dirac statistics. We can use eq. (51) and calculate their Fermi energy ($g = 2$, $N/V = Z/a^3 = n_e$):

$$\varepsilon_F = \frac{(3\pi^2 Z)^{2/3}\, \hbar^2}{2a^2 m}. \tag{53}$$

Since the *dimensions* of a unit cell of a crystal approximately coincide with atomic dimensions,

---

* This approach is certainly a very crude model. Many properties of metals, and especially those (as we have said above) caused by the application of a magnetic field, cannot be explained in terms of this model. However, the properties to be discussed below are not too sensitive to subtleties. They are determined above all by the *quantum properties of the gas of electrons.*

185

it is convenient to express a in units of the Bohr radius: $a = (\hbar^2/me^2)\,\beta$ (β is a dimensionless parameter equal to several units). As a result,

$$\varepsilon_{\mathrm{F}} = \frac{(3\pi^2 Z)^{2/3}}{\beta^2}\,\frac{me^4}{2\hbar^2}\,, \qquad (54)$$

$me^4/2\hbar^2 \approx 13.53$ eV (see Chapter 1) and we find that the Fermi energy is a quantity on the atomic scale. The adjective "atomic" is usually used as a synonym of "microscopic", "tiny" (although not in the expression "atomic explosion"). But here we have written the expression for ε_{F} above in such a form to emphasize that the Fermi energy is very high. If eq. (54) is used to calculate the Fermi temperature $T_{\mathrm{F}} = \varepsilon_{\mathrm{F}}/k_{\mathrm{B}}$, the result will certainly look suspicious: the degeneracy temperature of the electron gas in a metal is of the order of 10^5 K. This means that the electron gas in a metal is always very degenerate at room temperature and even up to the melting point of the metal, that is, any slab of metal is a natural "store" of a degenerate "quantum" gas with all its peculiarities and unusual features.

All our knowledge about the physics of the metal state is based on the concept of the electron gas as a degenerate gas obeying the Fermi-Dirac statistics. By assuming that the electron gas is a degenerate Fermi gas, it is possible to explain a variety of properties of metals (temperature dependence of electrical resistance, as well as thermoelectric, magnetic, and galvano-magnetic phenomena, etc.).

As early as the end of the nineteenth century it was known that, in contrast to dielectrics, metals contain free electrons. It was clear that

the number of electrons is very great: one or two electrons per unit cell of the crystal. Otherwise it would be impossible to explain the high electrical conductivity of metals and their lustre, that is, their ability to reflect light.

Still, there was a very simple experimental fact that spoiled the existing self-consistent picture. If one relied on the formulas of classical statistics, the specific heat of a metal must be higher than that of a dielectric; and not by a few percent—calculations indicated that the difference was by a factor of 1.5*. This difference could not be missed in experiments. However, the measured specific heat of metals stubbornly coincided with that of dielectrics. Free electrons could not be detected. By virtue of its simplicity, the problem was of fundamental importance. It was referred to as the "specific heat paradox". Everything becomes clear and the paradox is resolved naturally if the properties of a degenerate Fermi gas are taken into account. The temperature in the experiment is much lower than the Fermi temperature T_F, and under these conditions the electronic component of specific heat is simply much lower than the classical specific heat of the remaining "part" of the metal, that is, of its lattice.

But could the electronic component of the specific heat of the metal be measured, this confirming directly the existence of the degenerate electron gas? It would be meaningless to try and conduct experiments at temperatures much higher than the degeneracy temperature: almost

---

* The specific heat of the lattice is $3R$, and if there is one electron per atom, the specific heat of classical electrons must be $(3/2) R$ (see Chapters 2 and 3).

all metals would not only melt but simply evaporate at such a temperature. There is another method: to "go down" to very low temperatures.

This might seem silly at first glance. Electronic specific heat would be far too low. But low compared with what? Well, compared with its classical value. But what do we care how much lower the electronic specific heat is compared with its classical value? The only important thing is that it be higher than the specific heat of the crystal lattice, because in fact we measure the specific heat of the metal as a whole, that is, as the sum of the specific heat of the ionic lattice and that of the electron gas. However, the quantum effects described in the preceding section reduce the specific heat of the lattice much faster ($\sim T^3$) than the electron specific heat ($\sim T$) as the temperature approaches absolute zero. The two contributions become equal at a temperature of a few degrees. Figure 32 shows specific heat divided by absolute temperature as a function of squared absolute temperature for $T \ll \Theta$. The dashed line corresponds to a dielectric, while the solid line, to a metal.

This is not the Gedanken experiment. The electron specific heat of metals has been measured and the corresponding data are given in the available handbooks. These data are essential for understanding the nature of the metal state ... .

In order to emphasize the surprising properties of a degenerate electron gas, let us state the following fact that goes against our intuition (based on classical concepts): *a degenerate electron gas is more ideal the higher its density*, that is, the smaller the spacing between electrons. Although this statement may seem absolutely mean-
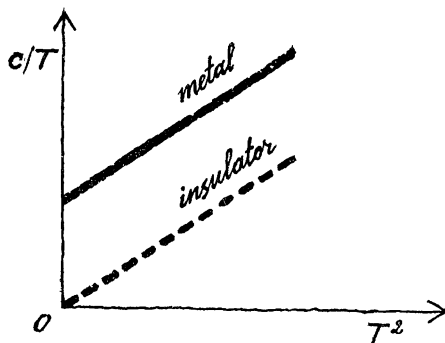
Fig. 32

ingless, let us analyze it and confirm its validity.

Indeed, a gas is the more ideal the lower the interaction energy between particles, U_{int}, when compared with their kinetic energy, E.* We have already mentioned that kinetic energy equals $(3/5)\ \varepsilon_{\text{F}}$; hence, it grows proportionally to the electron density n_e to the power of $2/3$ (E is proportional to $n_e^{2/3}$). Electrons interact according to Coulomb's law. The energy of interaction between two electrons is inversely proportional to the distance between them, that is, grows proportionally to the density to the power of $1/3$ (U_{int} is proportional to $n_e^{1/3}$). Naturally, as the density increases, the kinetic energy overtakes the potential energy, which is exactly what we have set out to prove. This characteristic of a degenerate gas offers a theoretical physicist a reliable method of analyzing the properties of electrons (the high-density approximation).

————————

* We mean estimates of both these energies per particle.

Chapter 5

Electrons and Phonons

Are electrons in a metal free?

There is an old joke about a newsman given to pompous phrases. He is reporting on the opening of a power plant. His account goes like this: "Mr. So-and-So (the name of a big boss) has plugged in the knife switch, and the electric current slowly began to flow along the mains ... ." The word "slowly" has been added to emphasize the solemnity of the occasion. Well, it is solemn all right, but grossly inaccurate. The first current signal will reach the consumer at a tremendous velocity, very near to the speed of light, but gradually (although also very rapidly) a steady-state current mode will set in. And if the current demand does not vary, the current in the circuit will remain constant.* However, the electric current is a flux of electrons. And what is the velocity with which electrons move in this flux?

Our question is not really valid, because we know that electrons in metals move even in a zero current. They do not rest even at absolute zero. This was described in detail in the preced-

---

* For the sake of simplicity, we assume that the power plant produces dc, not ac, current.
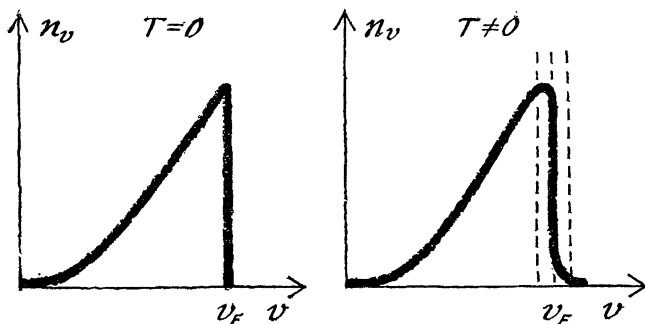
Fig. 33

ing chapters. We remind the reader that electrons in a metal move at different velocities, but that the bulk of the electrons moves at a velocity (denoted by v_F) approximately equal to 10^8 cm/s. This number can be obtained by using the formula $mv_F^2/2 = \varepsilon_F$ (see eqs. (53) and (54)). Figure 33 schematically shows the velocity distribution of electrons at absolute zero and at a temperature above zero.

Electrons thus move even when there is no current. This motion is, however, absolutely chaotic, resembling the random running-around of people in a mob. When current flows in a metal, the motion of the electrons resembles the random motion of people in a mob drifting slowly along the street. This time the word "slowly" is not added for solemnity but is meant to emphasize that the velocity of the advance of the mob, that is, the velocity of its ordered motion, is much lower than that at which people in a mob run around, that is, the velocity of random motion. Let us ascertain

first of all that our model agrees with reality. It is first necessary to estimate the mean velocity of electrons (we know the velocity of random motion). The current density is the quantity of electricity flowing each second across one square centimetre of the conductor, so that according to the definition

$$j = e n_e \bar{v}; \qquad (55)$$

$e = 4.6 \times 10^{-10}$ cgs esu, $n_e \approx Z/a^3 \approx 10^{22}$-$10^{23}$ cm^{-1}. We assume that current with density 1 A/cm^2 is flowing in a metal. Unfortunately, we cannot simply divide unity by the product $e n_e$; first, amperes have to be transformed into the electrostatic units we need; 1 A $= 3 \times 10^9$ cgs esu. We substitute these values and find, to our surprise, that $\bar{v} \leqslant 10^{-3}$ cm/s. As you see, electrons indeed move slowly, not only compared with the velocity of their random motion but even by our human everyday scale (the newsman whose pompous style opened this chapter was not, after all, as wrong as it seemed).

The velocity of the ordered motion of electrons is so low that we would be justified in wondering whether our assumption of treating electrons in a metal as free holds water. And what is the limitation imposed on this assumption?

What is the behaviour of a truly free electron? Current flows through the metal if a potential difference is applied to the ends of the conductor; this potential difference produces an electric field in the metal. We assume that this field E equals 1 V/cm. Therefore, force eE is applied to each electron, and the acceleration $w = eE/m$. It can be readily found that an electron travelling at this acceleration over a distance of 1 cm

gains a velocity of about 10^8 cm/s. This is nothing like the drift of electrons with velocity $\bar{v}$. The electrons are therefore far from being free. Their behaviour in a metal differs strikingly from that of truly free particles.

Ohm's law

The ability of a metal to conduct electric current became very early a subject of physicists' scrutiny. It was established as early as the beginning of the nineteenth century that the potential difference applied to a metal and the current produced by this field in a conductor are simply proportional to each other (*Ohm's law*); we shall present this law in the differential form:

$$j = \sigma E. \tag{56}$$

This relation is referred to as *Ohm's law in differential form*.

The proportionality factor between current density j and electric field strength E is electrical conductivity σ, while $\rho = 1/\sigma$ is the resistivity of the metal. In the usual notation, $R = \rho L/S$, where L and S are the length and the cross-sectional area of the conductor.

We readily find from the expression for the current density (eq. 56) that the mean velocity of electrons in the conductor is proportional to the force applied to the electrons:

$$\bar{v} = \frac{\sigma}{e^2 n_e} \, eE. \tag{57}$$

The quantity $u = \sigma/e^2 n_e$ is the velocity of a particle subjected to a unit force; it has been given a special name, *mobility*. Mobility character-

izes the force of resistance of the medium through which the electrons move. In the absence of this force electrons would be accelerated by the electric field and could not move at a constant mean velocity. *Mobility is a measure of the non-freedom of electrons.*

Equation (57) implies the similarity of the passage of electric current in a conductor and the flow of liquid in a pipe. A flow of liquid has to be sustained (usually by a pressure drop along the pipe) because of the friction of the liquid at the pipe walls. Not only those molecules of the liquid that directly impinge on the pipe walls are slowed down. There exists a mechanism of momentum transfer from rapidly moving parts of the liquid to its slowly moving parts. This is *viscosity*, a property that is explained on the basis of molecular kinetic concepts of the liquid state.

Electrons in a conductor also experience friction. However, not only walls exert friction on electrons. It would be better to say "it is not so much walls that exert friction on electrons", because friction at the walls is important only in very thin specimens (this will be discussed later). There are factors inside the metal that brake down the motion of electrons. The contents of the present chapter is mostly concerned with these factors.

It is customary not to express mobility as a function of electrical conductivity, but electrical conductivity as a function of mobility:

$$\sigma = n_e e^2 u. \tag{58}$$

The reader may feel puzzled: Why express a characteristic of a metal (electrical conductivity σ) via two other characteristics, n_e and u? Each

of these characteristics determines essentially different, or should we say independent, characteristics of the metal state. The number of free electrons per unit volume, n_e, is for all practical purposes independent of temperature and depends to only a slight degree on the state of the specimen of the given metal (whether this is a single crystal or a polycrystalline specimen, whether the specimen is very pure or contains minute amounts of impurities). Mobility u may vary with temperature by a factor of hundreds or thousands; the removal of impurities from the metal increases mobility. Both characteristics, namely, the number of electrons per unit volume n_e and mobility u, can be measured independently.

The Hall effect. Measuring mobility

One of the most acceptable methods of measuring the number of electrons in a conductor is a technique based on the Hall effect, that is, on measuring the potential difference in a plate specimen placed in a magnetic field when an electric current is passed through the specimen (Fig. 34).

The cause of the Hall effect is very simple. The force applied to an electron by the electric and magnetic fields is the Lorentz force $e \left(\mathbf{E} + \right.$ $\left. + \frac{1}{c} [\mathbf{v}\mathbf{H}] \right)$, which makes the electron move in a direction normal to vectors $\mathbf{E}$ and $\mathbf{H}$. In this direction the electric circuit is not closed and, therefore, the redistribution of charges produces friction, that is, an electric field directed along the y-axis (the coordinate axes are shown in Fig. 34).
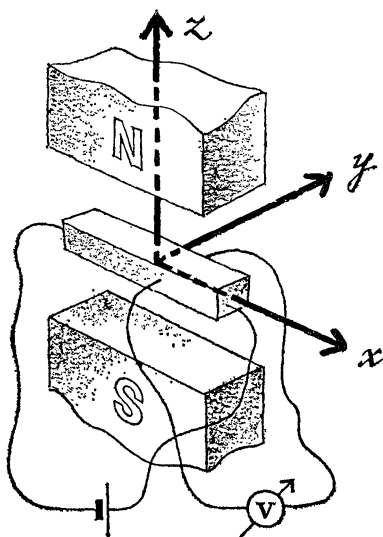
Fig. 34

We easily obtain after simple transformations that

$$E_y = \frac{1}{n_e e c}\, jH. \tag{59}$$

The electron's charge and the speed of light are known, while j, H, and E_y are measured directly; therefore, eq. (59) enables us to find the number of electrons per unit volume of a conductor. Such measurements demonstrate that a good metal has $n_e \approx 10^{23}$ cm^{-3}. This confirms our estimate: each atom of the metal releases one or even several electrons that freely move in the crystal. It might be thought that this is another proof of

the validity of our free electron model, especially if we take into account that this figure is in agreement with other measurements, such as measurements of electronic specific heat.

The knowledge of the number of electrons per unit volume readily yields the electron mobility in a metal via eq. (58). Resistivity ρ of a metal can be found in a handbook. As a rule, a handbook gives ρ at room temperature. After the necessary calculations, we arrive at a figure; for example, we find that at room temperature $u = 2 \times 10^{13}$ s/g for copper. Is this high? It is very low if we compare it with the electron mobility in copper at very low temperatures. Close to absolute zero, mobility is hundreds or even thousands of times higher. If, however, we compare it with the electron mobility in other conductors, then it is rather high: copper has one of the highest electron mobilities.

Mean free time. Mean free path

With mobility known, we easily estimate two other very important characteristics of electrons in a metal, those included in the heading of this section. The dimension of mobility u is s/g. Therefore, the quantity $\tau = mu$, where m is the electron's mass, includes the dimension of time, s. Different names are used for the quantity τ (a not infrequent situation). The most popular terms are *time constant* and *mean free time*. These terms have come from different fields of physics: the first from mechanics and the second from the molecular kinetic theory of gases.

An analysis of the motion of a particle with nonzero friction shows that immediately after

an external force is switched on (say, an electric field is switched on) the particle moves at acceleration, and its velocity increases; at the same time, the friction force proportional to velocity is also increased. After a period approximately equal to $\tau = mu$, the force of friction completely balances out the external accelerating force, and the particle begins to move at a constant velocity. The time constant thus characterizes the build-up time for the steady-state mode of motion, that is, the motion at a constant velocity.

Let us estimate the time constant. In our example, $\tau \approx 2 \times 10^{-14}$ s. This is fantastically small. It is impossible to detect the period of free motion in dc experiments. In practice Ohm's law takes over in a conductor immediately after the conductor is switched into an electric circuit.*

Despite the fact that the time constant τ of electrons is so fantastically short, it is nevertheless possible to observe the period of free motion of electrons (their drift) and even to measure τ. This requires a study of the behaviour of metals in high-frequency electric fields.

Microscopically, the motion of electrons with friction is a sequence of periods of motion under the action of an accelerating force and collisions in which the particle dissipates the energy that it gained when accelerated. In this approach, τ is the mean time between two collisions, and the product $v_F \tau = l$ is the mean free path, or

---

* In reality the steady-state mode is not immediately established on closing the circuit. The relaxation time is much longer than τ but is determined by the inertial properties of the whole circuit and not of the electrons in the metal.

its range, that is, the mean distance covered by a particle between two collisions.

This means that the nature of electrical resistance in conductors can be understood by explaining why electrons in a crystal have a finite mean free path and calculating its numerical value.

Quantum mechanics has explained the "liberation" of an electron from inside the atom. This has led to the conclusion that in an ideal crystal, that is, in an infinite crystal with no defects and in which each ion is immovably fixed to its site, an electron will be offered no resistance to its motion and will ignore the surrounding ions. By accelerating under the action of the applied field, it will increase its energy until it comes to the boundary of the specimen.

To conclude this short section, we shall give several formulas that express electrical conductivity in terms of microscopic quantities: the electron's charge e, its mass m, the mean free path l, etc. Equation (58) and the definition of the time constant τ yield

$$\sigma = \frac{n_e e^2 \tau}{m} \, . \tag{60}$$

If both the numerator and the denominator are multiplied by v_F, then the expression for electrical conductivity is recast in the form:

$$\sigma = \frac{n_e e^2 l}{m v_F} \, . \tag{61}$$

Since $m v_F = p_F$, eq. (61) can be rewritten in the form:

$$\sigma = \frac{n_e e^2 l}{p_F} \, ; \tag{62}$$

$4\pi p_F^2 = S_F$ is the surface area of a sphere of radius p_F. Therefore, eq. (62) can be rewritten as follows:

$$\sigma = \frac{e^2 S_F l}{12\pi^3 \hbar^3}, \quad S_F = 4\pi \left(3\pi^2 n_e\right)^{2/3} \hbar^2. \quad (63)$$

The significance of this formula lies in the fact that it clearly shows that a charge in a metal is transferred by electrons located at the Fermi surface, because only these electrons can absorb energy from the electric field; indeed directly above the energy levels of these electrons there are allowed states free from other electrons. Besides, this formula is quite "up-to-date"; it does not change if we take into account the complicated dynamics of conduction electrons when the Fermi surface is not spherical. But in this case it is not possible to express S_F via the number of electrons per unit volume, n_e (see eq. (63)).

We have seen that in many metals the Fermi surfaces are quite complex: they may be branched, or have numerous "arms", "pockets", and "bottlenecks" (this picturesque imagery has become quite acceptable for serious articles on the physics of metals). In calculating the electrical conductivity factor, it is necessary to take into account the complexities of the Fermi surfaces since calculations by well-familiar formulas valid for metals with spherical Fermi surfaces may lead to serious mistakes*.

---

* It may help to make this less abstract, not so remote from real metals, if the reader keeps in mind that all alkali metals (with an exception of Li) have spherical Fermi surfaces.

Electrons transfer heat.
The Wiedemann-Franz law

A short digression. Let us imagine a running competition with very unusual rules that require that a participant must carry away from the start, and as far as he can, a hat on his head and a flag in his hand. His task is quite complicated because no runner is told the route and so he has to choose it by random signs. Moreover, his way is full of people who are allowed to knock the hat off his head and to snatch the flag out of his hand. We evaluate the success of each participant by how far he manages to carry the hat and/or flag along the correct route. It is clear that once the runner loses the right path, it is of no significance whether he manages to hold on to his flag and hat or not. In some cases the runners following the right route will carry the flag longer and sometimes they will carry the hat longer. This depends on the behaviour of people along the path and the capabilities of the runner.

Now let us give a definition that will help us again to pick up the thread of our story. Let us define the mean distance over which the runners manage to carry the hat as the mean free path with respect to the loss of hat, and the mean distance over which the flag is carried as the mean free path with respect to the loss of the flag. Obviously, these two mean free paths are different in the general case. They become identical only if the runners simply wander off the right path.

This story must show the reader that the mean free paths are not necessarily identical. To be more exact, we must add after the words "mean

free path" the qualification "with respect to a loss" and mention what exactly is lost in the collision (not by a runner but by an electron!).

Electric current is the transfer of electric charge. Electrons cannot lose their charge and will fail to transfer their charge only if they lose the path.

The motion of electrons can transfer not only the electric charge but heat as well, if the electrons move from the hotter end of the specimen to the colder end. Electrons participate in heat conduction. As a rule, the heat conduction of metals is higher than that of dielectrics. The reason is again the participation of electrons in heat transfer.

There is nothing surprising about this. In any gas particles transfer heat from the hotter wall to the colder wall. A metal is a container filled with "gases", that is, phonons and electrons. The calculation of the electronic heat conductivity $\varkappa$, carried out with assumptions just as general as the derivation of eq. (63), yields the following result:

$$\varkappa = \frac{2\pi^2 k_B^2}{9\ (2\pi\hbar)^3}\ T S_F l_h.\tag{64}$$

The mean free path l_h is given a subscript to underline that this is the mean free path with respect to the loss of heat by electrons. When l_h is calculated, it is necessary to take into account how much energy is lost by an electron in each collision. The mean free path of the electron (without subscript) in the expression for electrical conductivity is the mean free path with respect to the loss of the direction of motion. In order to avoid complications in the notation,

we shall simply refer to the mean free path. In the general case, l and l_h are different and the ratio

$$\frac{\varkappa}{\sigma T} = \frac{\pi^2 k_B^2}{3e^2} \frac{l_h}{l} \tag{65}$$

depends on the type of metal and is a function of temperature. At ordinary temperatures, however (room temperature and higher), l_h is equal to l to a very high degree of accuracy, and the ratio $\varkappa/\sigma T$ is a constant quantity dependent neither on temperature nor on the type of metal. It is equal to $\pi^2 k_B^2/3e^2 = 2.45 \times 10^{-8}$ W·Ω/K.

The conclusion on the constancy of the ratio $\varkappa/\sigma T$ for different metals was obtained experimentally and called the *Wiedemann-Franz law*. The Wiedemann-Franz law is an excellent proof of our concepts concerning the internal structure of metals. Moreover, the exact value of the constant in the Wiedemann-Franz law is a direct proof of the degeneracy of the electron gas. If electrons were described by classical statistics, the constant would be different ... .

Investigations of electrical and heat conduction in metals at low temperatures demonstrated that the Wiedemann-Franz law is far from being universal. Of course, this does not reduce our confidence in the main ideas on the internal structure of metals. Observed departures from the Wiedemann-Franz law give additional evidence of the correctness of our theory, because l_h is not identical to l. Experiment shows when this difference is most important, while theory pinpoints the cause of the discrepancy ... . This will be the topic of one of the sections to follow.

What is the stuff that electrons rub against?

When we were speaking about the mean free path or the mean free time of electrons, the term "mean" was defined in two ways: first, "averaging" is carried out over all electrons (one electron is scattered differently from another) and, second, the obstacles on which electrons are scattered are also averaged.

There can be no ideal crystals in nature. Each departure from perfect periodicity is, for the electron, and obstacle. An example of such an obstacle is an impurity atom that was accidentally incorporated into the crystal, or a boundary of a crystallite, or a vacancy, that is, a site in the lattice in which there should have been an ion if the order were ideal, etc.* Real crystals contain many and extremely varied defects. At ordinary (room) temperature the main cause of the scattering of electrons, and with it of the electrical resistance of metals, is not the presence of defects in the crystal lattice but the thermal motion of ions. If the temperature is very low, that is, close to absolute zero, the thermal motion is almost suppressed and the defects that have been discussed above are clearly "seen".

Let us divide the collisions of electrons into two classes. We shall treat separately collisions that involve departures from (defects in) the periodicity of the crystal lattice and those that are caused by the thermal motion. Electron-

---

* The "wave language" better conveys the feeling that any departure from periodicity is an obstacle. See p. 169.

electron collisions will be subsumed under the second class.

As a result of thermal motion, electrons would have a finite mean free path even if we could create an ideal infinite crystal, although this mean free path would increase infinitely with decreasing temperature.

We shall call the mean free path in such an ideal crystal the ideal mean free path l_{id}, and the resistivity of the ideal crystal, the ideal resistivity ρ_{id}.

At absolute zero electrons in a real crystal collide only with defects in periodicity. The mean free path limited by such collisions is called the residual mean free path l_r and the corresponding resistivity, the residual resistivity ρ_r. The resistivity is called residual because a residue of resistance *is retained* even at absolute zero.

There is a simple rule which states that the total resistivity ρ of a real metal is the sum of the ideal and residual resistivities (Mathiessen's rule):

$$\rho = \rho_{id} + \rho_r. \tag{66}$$

Experiments give the following results: two specimens of the same metal have very similar curves of the temperature dependence on resistivity (Fig. 35), one curve being shifted with respect to another by $\rho_r^{(1)} - \rho_r^{(2)}$ (the superscripts are to enumerate the specimens). Rule (66) was derived by analyzing experimental data.

According to Mathiessen's rule, there is no need to try to produce an ideal specimen to find out the ideal resistivity of a metal (an ideal

ρ, ohm·m ×10⁻⁸

Cu + 3.32% Ni

Cu + 2.16% Ni

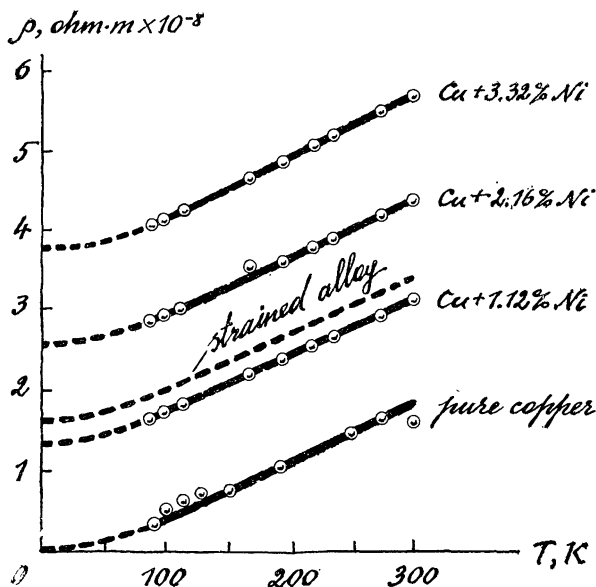strained alloy

Cu + 1.12% Ni

pure copper

T, K

Fig. 35

specimen is simply impossible, even if only because no specimen can be infinitely large). We have to plot a curve showing the resistivity of an actual specimen as a function of temperature, $\rho\,(T)$, and to extrapolate this curve to absolute zero. This will give ρ_r for the specimen ($\rho_r = \rho\,(T \to 0)$); subtraction of ρ_r from $\rho\,(T)$ yields $\rho_{1d}\,(T)$ of the metal. As a rule, experimenters carry out this procedure with several specimens and are satisfied if the result is independent of the specimen used, within the accuracy of the

206

experiment. Note that the method of processing the experimental data as given above assumes that ρ_r is independent of temperature, and $\rho_{1d}(T)$ is independent of the state of the specimen. Experiments confirm this assumption rather poorly. Deviations are observed, called deviations from Mathiessen's rule. In recent years numerous experimental and theoretical papers have been devoted to studying these deviations.

What is the nature of Mathiessen's rule? Why are the resistivities, caused by different scattering mechanisms, additive?

The mean free path and the mean free time have a simple probabilistic interpretation: $w = 1/\tau$ is the probability of collision (scattering) of the "mean" electron per unit time, and $1/l$ is the probability of collision per unit free path. If there are several possible causes of scattering in the crystal, the probability of collision is the sum of the probabilities of collisions with each obstacle*:

$$w = \sum_j w_j. \tag{67}$$

Subscript j enumerates the mechanisms of scattering. If we separate all the factors involved in scattering in an ideal crystal, we obtain

$$w = w_r + w_{1d}.$$

A comparison of w_r, that is, the residual mean free path, with w_{1d}, that is, the ideal mean free

---

* In mathematics this formula is called the theorem of the addition of probabilities. The condition for its validity is the independence of events.

path, yields

$$\frac{1}{l} = \frac{1}{l_r} + \frac{1}{l_{1d}}.$$ (68)

The last relation is equivalent to Mathiessen's rule, since the resistivity $\rho = 1/\sigma$, and the proportionality factor between ρ and $1/l$ is independent of the mechanism of scattering (see eqs. (60)-(63)). Equations (66) and (68) are extremely important because they enable us to distinguish between various mechanisms of electrical resistance and study each one of them independently.

The largest contribution (this figurative expression means: the largest term in the sum of probabilities) to the *residual* resistivity is made by the scattering on various impurities which are present in any crystal. The mean free path l_{imp} due to scattering on impurities is easily estimated by assuming an impurity atom to be a sphere with a radius close to the dimension of a unit cell. Then of course we obtain

$$\frac{1}{l_{imp}} \approx N_{imp} a^2,$$ (69)

where N_{imp} is the number of impurity atoms. In order to replace this approximate equality by an exact one, we have to calculate, using our knowledge of the forces acting between the impurity atoms and electrons, as well as the laws governing the motion of electrons (quantum mechanics), a quantity which is called the effective scattering cross section. This quantity is substituted for the factor a^2 in eq. (69). The exact calculation does not alter l_{imp} very much but does make it possible to find out which of the impurities affects the resistivity of the metal.

Collisions with impurity atoms, as well as collisions with other departures from the periodicity of atoms in crystals, leave the energy of electrons practically unaltered. But the direction of motion of electrons is changed—they leave the path they are following. Impurity atoms are heavy and electrons are light. Consequently, if impurities constitute the main cause of scattering, the mean free path with respect to the loss of heat, l_h, is indistinguishable from the mean free path with respect to the loss of direction, that is, from the mean free path l. Once the path is lost, neither charge nor heat excess can be transferred. Therefore, the Wiedemann-Franz law is valid if the mean free path equals l_{imp}. From what we have said earlier it is clear that this will happen at very low temperatures.

What will the situation be at higher temperatures, such as room temperature? In order to answer this question, we have to clarify the effect of the thermal motion of metal ions on resistivity. This topic is so important that we shall devote a special section to it.

Electrons collide with phonons

It is customary to begin a description of the role played by the thermal motion of ions in the resistivity of metals with the words "In electron-phonon collisions ...".

We must, however, define the term "collision" more accurately. We automatically assume that "collision" means that one object impinges on another, while in the case of electron-phonon

interactions the basic collision processes are*:
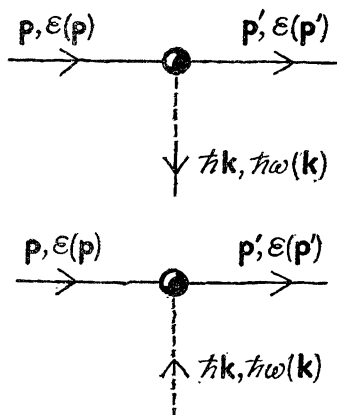
Solid lines symbolize electrons, and dashed lines, phonons; the lines are labelled by the momenta and energies of the electrons and phonons (the primed quantities refer to quantities after the collision). In the second process it is indeed a collision of an electron and a phonon. The outcome, alas, is the death of the phonon. In the first process there is in fact no collision at all, but the electron creates a phonon. Let us check first of all whether the processes shown are allowed by the laws of momentum and energy conservation. We readily find that within the accuracy of notation the laws are written identically for both cases:

$$\varepsilon\,(\mathbf{p}) + \hbar\omega = \varepsilon\,(\mathbf{p} + \hbar\mathbf{k}). \qquad (70)$$

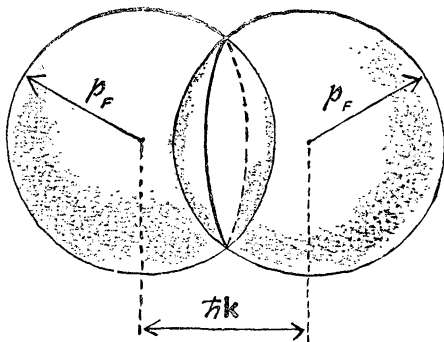* Ordinary collision processes also occur, but they are less probable than those shown in the drawing.

210

Fig. 36

We are interested in electrons whose energy is equal to the Fermi energy, but for which $\varepsilon_F \gg \hbar\omega_{max} = k_B\Theta$. Indeed, $\Theta \approx 10^2$ K, $T_F \approx 10^5$ K. Hence, we can simply ignore $\hbar\omega$ compared with ε_F and rewrite the conditions of creation (annihilation) of a phonon with momentum $\hbar\mathbf{k}$ in the form:

$$\varepsilon_F(\mathbf{p}) = \varepsilon_F(\mathbf{p} + \hbar\mathbf{k}). \qquad (71)$$

This shows that the process is "allowed" if the Fermi surface $\varepsilon(\mathbf{p}) = \varepsilon_F$ and its analogue (shifted by $-\hbar\mathbf{k}$) $\varepsilon(\mathbf{p} + \hbar\mathbf{k}) = \varepsilon_F$ *intersect* (Fig. 36). If $\hbar k < 2p_F$, condition (71) is satisfied, that is, the creation and annihilation of a phonon are allowed*. There is no limit to the number of phonons with $\hbar k < 2p_F$: the Fermi surface occupies a predominant part of the first Brillouin zone.

---

* The fact that the process of creation and absorption of phonons by electrons is forbidden for $\hbar k \geqslant 2p_F$ results in a curious feature in the phonon spectrum of metals (the Migdal-Kohn anomaly).

The collision of an electron with an acoustic phonon is an interesting illustration of the correspondence principle. Let us formally have $\hbar$ in eq. (71) tend to zero. Then eq. (71) yields $\mathbf{v_F k} = 0$, that is, electrons participating in the process of creation (annihilation) of a phonon have velocities perpendicular to the wave vector. True, once we switch to the classical description, we should speak not about the creation (annihilation) of a phonon but about emission (absorption) of a sound wave. But this is not all. Also, as the high Fermi energy is cancelled out, it is better to retain $\hbar\omega$. As a result, we obtain

$$\mathbf{kv_F} = \omega \quad \text{or} \quad \cos\Theta = \frac{s}{v_F}, \tag{72}$$

where Θ is the angle between the electron's velocity and the wave vector. The emission (absorption) of sound waves is possible if $v_F > s$. But this condition is always satisfied: $v_F \approx 10^8$ cm/s and $s \approx 10^5$ cm/s. A comparison of conditions (70) and (72) shows that the creation (annihilation) of a phonon is a quantum analogue of the Cerenkov emission of sound.

Agreement with the conservation laws is not enough for a collision to take place. There must also be a *real* interaction between colliding particles (recall the arguments about the anharmonicities which make the interaction between phonons possible). What is the nature of the electron-phonon interaction? A phonon is a quantum of vibrations of the metal ions. Electrons move in the field of ions. When ions are displaced from their equilibrium positions, the electron-ion interaction energy is changed. The change in the electron's energy due to the displacement of ions

is precisely the electron-phonon interaction energy.

Like all bosons, phonons have a spectacular property: the intensity of their creation in any state is the higher the greater the number of bosons in that state.* Bosons are motivated by the "herd instinct": "Go where everybody else goes." However, bosons can be created into a vacuum as well.

A metal has no phonons at absolute zero. It seems logical to expect that resistance caused by collisions with phonons will not disappear as the temperature tends to zero. An electron accelerated by an electric field cannot indeed *absorb* a phonon because there are no phonons; however, electrons seem to be able to create phonons and thus may be slowed down. This line of argument is flawless as long as we are dealing with a single electron, but once we turn to a huge ensemble of electrons the situation becomes much more complicated.

Calculations show that the part of resistance which originates from collisions with phonons vanishes as the absolute temperature tends to zero. This contradiction may undermine our belief in the correctness of our qualitative analysis of the mechanism of electrical resistance. However, if we follow through all the calculations which are found in any textbook on the electron theory of metals, we find that the mechanism of resistance formation is described correctly: each collision is accompanied by the emission or

---

* This property of bosons, discovered by Albert Einstein, is fundamental for the functioning of lasers and masers.

213

absorption of a phonon.* What then is the explanation?

The reader already knows that the motion of a collective (ensemble) of electrons is described by the mean velocity, which is very small. A "mean electron" can neither emit sound waves nor generate phonons. But it can be scattered on vibrating ions. A "mean electron" is a very complex concept. Indeed, some electrons moving along an electric field are accelerated by the field, while those moving against the field are slowed down. The first are slowed down and thus create phonons, and the second absorb phonons and are therefore accelerated. Otherwise they would be left behind by their kinsmen. There exists a well-developed mathematical technique that makes it possible to understand all these complicated processes and consistently pass from describing individual events of the emission and absorption of sound waves by electrons to such macroscopic concepts as mobility, mean velocity, and electrical conductivity. These problems are within the realm of a special branch of theoretical physics, *physical kinetics*.

The following phenomenon demonstrates that our concepts of mean velocity have a profound physical meaning: once the conditions are created in which the mean velocity of the directional drift of electrons exceeds the speed of sound, the

---

* The linear relation between current and field (eq. 56) which we are studying here is the first term of the expansion of j in powers of E. The creation of phonons by electrons at $T=0$ results in the dependence of conductivity on the field and is not covered by Ohm's law. Strictly speaking, $\sigma = \lim_{E \to 0} (j/E)$.

resistivity of the conductor sharply increases—"mean electrons" begin to create phonons.

Let us return to the mechanism of ideal electrical resistance. The scattering of a "mean electron" is described more simply without resorting to phonons, although we shall have to appeal to phonons again to explain the nature of deviations from the Wiedemann-Franz law. The approach we have chosen enables us not only to understand the nature of electrical resistance but also estimate the mean free path.

Thermal motion violates the rigorous periodicity of the lattice, so that scattering results from random vibrations of lattice ions; the probability of scattering w (we recall that w is inversely proportional to the mean free path: $w = v_F/l_{ph}$) is the higher the larger the mean squared amplitude of ionic vibrations, $\overline{x^2}$.

The mean free path l_{ph} is given by a simple formula:

$$\frac{1}{l_{ph}} = N'\overline{x^2}, \tag{73}$$

where N' is the number of ions per unit volume. This formula indicates that the scattering probability is proportional to the area "occupied" by a vibrating ion in the plane perpendicular to the electron's velocity.*

---

* Equation (73) somewhat masks the resistivity paradox at $T = 0$, but does not resolve it. At $T = 0$ the amplitude of vibrations of ions is nonzero and we again have to pose the same question: Are electrons scattered on zero-point vibrations? The answer is "No". The fact is that zero-point vibrations are correlated or, to use a vaguer term, uniform. Once this uniformity is violated, zero-point vibrations are reflected in the value of resistivity at $T = 0$. Uniformity is automatical-

Numerical multipliers will be dropped in the formulas both in this chapter and hereafter. In some cases these factors may considerably change the estimate of a parameter. The temperature dependence will be given correctly. The mean free path will be marked by a subscript "ph" since we shall speak essentially of the mean free path of electrons limited by the scattering on phonons.

Calculations show that

$$\frac{\overline{x^2}}{a^2} = \begin{cases} \dfrac{k_B T}{M s^2}, & T \gg \Theta, \\[2mm] \dfrac{k_B T^3}{M s^2 \Theta^2}, & T \ll \Theta. \end{cases} \qquad (74)$$

The last relation indicates that $\overline{x^2} \ll a^2$ at all temperatures. If electrons were scattered by randomly arranged atoms, the corresponding mean free path would be equal to $(N' a^2)^{-1*}$ (cf. formula (69)), while in fact it is much greater. Equation (62) and the relations given above readily lead to the value of the phonon component of resistivity:

$$\rho_{ph} \approx \begin{cases} \dfrac{a}{v_F} \dfrac{k_B \Theta}{M s^2} \dfrac{T}{\Theta} & \text{— correct!} \\[3mm] \dfrac{a}{v_F} \dfrac{k_B \Theta}{M s^2} \left(\dfrac{T}{\Theta}\right)^3 & \text{— incorrect!} \end{cases} \qquad (75)$$

ly violated if the specimen is not monoisotopic. Indeed, the nonuniformity of zero-point vibrations produces non-zero resistivity at $T = 0$ (this has been shown by I. Ya. Pomeranchuk). Isotopes cannot be treated as normal impurity atoms since all isotopes have identical electron shells and therefore scatter electrons in an identical manner.

* The mean free path of a particle in a liquid $l_{lq} \approx a$, where a is the interatomic spacing: $l_{lq} = (N' a^2)^{-1} = (a^2/a^3)^{-1} = a$.

We assumed, for the sake of simplicity, that $n_e = N'$ and $mv_F^2 = e^2/a$, with both equalities usually holding within an order of magnitude.

The first line in relation (75) is the classical expression for resistivity and is borne out by the exact theory, while the second line gives the quantum expression for ρ_{ph} and is *incorrect*: it must be multiplied by $(T/\Theta)^2$. The fact is that as temperature diminishes, not only the amplitude of the vibrations of ions but also the velocity of these vibrations decrease. Therefore each collision with a vibrating ion only slightly changes the motion of an electron, the collisions thus being ineffective. About $(\Theta/T)^2$ collisions are required before an electron "notices" that something is blocking its way (a single collision is sufficient only at $T > \Theta$). Hence,

$$\rho_{ph} \approx \frac{a}{v_F} \frac{k_B\Theta}{Ms^2} \left(\frac{T}{\Theta}\right)^5, \quad T \ll \Theta - \text{correct!}$$
(76)

This formula was derived by Bloch in 1929.

The low efficiency of collisions between electrons and vibrating ions is an important phenomenon. It constitutes the main cause of deviations from the Wiedemann-Franz law. Let us discuss it in "phonon terms". A collision changes the electron's energy by the phonon's energy, and its momentum by the phonon's momentum. If the increment in the electron's energy is denoted by $\Delta\varepsilon$, and the increment in its momentum by $\Delta\mathbf{p}$, then

$$\Delta\varepsilon = \hbar\omega, \quad \Delta\mathbf{p} = \hbar\mathbf{k},$$

where ω and $\mathbf{k}$ stand for the frequency and wave vector of the absorbed phonon. At low tempera-

tures $(T \ll \Theta)$ phonons in a solid usually have an energy close to $k_B T$, and momentum $\sim (\hbar/a)$ (T/Θ) (see Chapter 2). As a result, these are the phonons that are absorbed by electrons with the highest probability. Therefore, $\Delta\varepsilon \sim T$ and $\Delta p \approx (\hbar/a)(T/\Theta) \approx p_F(T/\Theta)$ since $p_F \approx \hbar/a$. A significant change in the direction of the motion of an electron means a change in its momentum by a quantity equal, by an order of magnitude, to p_F. An estimate of $\Delta\mathbf{p}$ shows that the change in momentum caused by an absorption of a phonon is very small in comparison with the Fermi momentum and confirms our statement that a great number of collisions is required to produce an appreciable deflection of momentum. An exact calculation shows that $(\Theta/T)^2$ collisions are required.

When studying heat conduction, we need to know how electrons dissipate heat. We have seen that each collision changes the electron's energy by about $k_B T$. Although electrons almost never stray from their path, they lose heat energy very easily.

In other words, the mean free path l_h with respect to heat losses is much shorter in collisions with phonons than the mean free path l with respect to the loss of direction. The ratio of these mean free paths, $l_h/l \approx (T/\Theta)^2$, is a measure of inefficiency of electron-phonon collisions at low temperatures.

We probably do not need to emphasize that the Wiedemann-Franz law must be violated in cases when the scattering on phonons is the main cause of scattering at low temperatures $(T \ll \ll \Theta)$. This conclusion is experimentally confirmed.

Most phonons at high temperatures ($T \gg \Theta$) have an energy equal to $k_B \Theta$, and momentum approximately equal to the Fermi momentum. Therefore each collision causes an electron to "stray off the path" completely. Evidently, $l_h \approx l$ and the Wiedemann-Franz law holds.

Electrons collide with one another

"Gas" is a word that brings to mind an image of thousand millions of particles moving in all directions with different velocities, colliding with one another, and in these collisions changing their directions of motion, that is, exchanging energy and momentum.

Electrons in metals also collide with one another. What is the contribution of electron-electron collisions to the resistivity of metals?

We posed a similar question when discussing the phonon heat conduction (see Chapter 2). And the answer is again similar: such collisions play an appreciable role if they are accompanied by Umklapp processes. But there is a distinction here, and a very substantial distinction. In interelectron collisions the Umklapp processes occur quite frequently since, as a rule, the Fermi surface is not far removed from the boundaries of the first Brillouin zone. Therefore the expression for the mean free path does not contain the characteristic (for phonons) exponential factor and is proportional to T^2:

$$\frac{1}{l_e} \approx \frac{p_F}{\hbar} \left(\frac{e^2}{\hbar v_F} \right)^2 \left(\frac{k_B T}{\varepsilon_F} \right)^2. \tag{77}$$

This formula (its derivation was given by L. Landau and I. Pomeranchuk) is not given

249

here to be used for estimating the electron-electron mean free path. The value obtained by formula (77) is not really very close to the experimental value. But we want to emphasize the fact that as temperature tends to absolute zero, the interelectron mean free path tends to infinity.

This may sound puzzling. Indeed, the number of electrons does not depend on the temperature. Why then do the electrons stop colliding? Well, they do not stop colliding, but the results of collisions cannot be detected. All states with energy below the Fermi energy are occupied both before and after the collision, and we cannot detect which of the electrons occupies which state: the electrons are *indistinguishable*. The indistinguishability of electrons is one of the basic principles of quantum mechanics. A corollary of this principle is an infinite interelectron mean free path at absolute zero.

Resistivity being proportional to $1/l$, that part of resistivity that originates by electron-electron collisions is proportional to squared temperature ($\rho_e \sim T^2$). Actually this term (we mean Mathiessen's rule) is more or less readily observable in transition metals (Pt, Fe), while in other metals it is lost against the background of other, larger terms.

A summary

At high temperatures ($T \gg \Theta$) the main factor producing the resistivity in metals is the scattering on thermal vibrations of ions; as a result, the total resistivity is given by relation (75) and the Wiedemann-Franz law holds. At low tempera-

tures $(T \ll \Theta)$ the resistivity is the sum of three terms:

$$\rho = \rho_{\text{imp}} + \rho_{\text{ph}} + \rho_{\text{e}}. \qquad (78)$$

The first term is independent of temperature, the second is proportional to T^5, and the third is proportional to T^2.

Of course, the true properties of metals not related to the preparation procedures of a given (random) specimen can be found by investigating the ideal part of resistivity: $\rho_{\text{id}} = \rho_{\text{ph}} + \rho_{\text{e}} + \ldots$. The ellipsis is meant to emphasize the fact that in principle there may be other factors contributing to resistivity.

As ρ_{id} decreases with temperature, it is logical to try and diminish ρ_{imp}. In the current state-of-the-art metals are produced so pure that in fact the mean free path is no longer than a microscopic parameter. In the purest metals it reaches, in the vicinity of absolute zero, several millimetres. At room temperature, though, the mean free path is very small, approximately 10^{-6} cm. One millionth of a centimetre is a very short distance on the human scale. In the world of atoms, though, 10^{-6} cm is not that small: it is approximately one hundred interatomic spacings. We find therefore that collisions do not restrict the freedom of electrons if this freedom is defined as the freedom of motion and if we recall that within atoms electrons do not move away from the nucleus by farther than a few angstroms.

Chapter 6

Magnons

Presumably, every child is surprised when first shown a magnet. But not every child is so moved by this event that he starts pondering over phenomena in the surrounding world, as was the case with Albert Einstein. But every one definitely retained the feeling of witnessing a miracle for quite a while. "A wonder of such nature I experienced as a child of 4 or 5 years, when my father showed me a compass. That this needle behaved in such a determined way did not at all fit into the nature of events, which could find place in the unconscious world of concepts (effect connected with direct 'touch')."*

Unfortunately, "the development of this thought world is in a certain sense a continuous flight from 'wonder'."**

We know that the action-at-a-distance, which so impressed the five-year old Albert, is caused by the magnetic field around a magnet (the needle of the compass is moved by the earth's mag-

---

  * *The Library of Living Philosophers. Albert Einstein: Philosopher-Scientist.* Ed. by P. A. Schilpp, Tudor Publishing Company, 1949.
  ** *Ibid.*

netic field); all characteristics of this field can be calculated once the power and shape of the magnet are known. But what is a magnet? Why are some bodies magnetic and others not? Mankind discovered magnetic properties so long ago that even the etymology of the word "magnet" is lost. But although the history of the useful applications of magnets is also quite long (the compass is the basis of navigation), the nature of magnetism was finally understood only rather recently (in the thirties of this century).

We shall try to explain why some bodies are magnets and others not by assuming that the main building blocks of matter (electrons, protons, and neutrons) are tiny magnets. Since the magnetic moments of protons and neutrons are lower than that of electrons by a factor of approximately one thousand, we are justified, at least at the initial stages of our study, in ignoring the magnetism of atomic nuclei; in fact the magnetic properties of iron are totally determined by the behaviour of its electrons.

Two types of atoms

In order to understand the magnetic properties of macroscopic bodies, that is, crystals, it is extremely important to know the structure of the atoms of specific materials. The quantum theory of atoms gives an unequivocal answer to the question of which atoms have nonzero magnetic moment, what its value is, and which atoms have zero magnetic moment. This answer is determined by the distribution of electrons over the energy levels of the atom. Figure 3 shows that the lower levels are occupied by paired electrons,

so that their net spin is zero, while the upper levels are occupied by unpaired electrons, so that their net spin is nonzero. This means that the magnetic moment also differs from zero. It is always necessary to use a good deal of energy (on the atomic scale) in order to change the magnetic moment of an atom. For example, a helium atom will be turned into a particle with nonzero magnetic moment if we "lift" one electron to the energy level that is separated from the ground level by a "distance" of 19.77 eV. So it is logical in what follows to consider only those atoms that have nonzero magnetic moments.

Two remarks which complicate the picture. First, the magnetic moment of an atom may differ from zero not only because of the electron spin but also because of the ordinary spatial (the more usual term is "orbital") motion of electrons.* Second, atoms with zero magnetic moment (referred to as diamagnetic) also have magnetic properties. A magnetic field induces in them a low magnetic moment directed against the magnetic field. This moment is so small that objects composed of diamagnetic atoms are said to possess weak magnetism, in contrast to objects composed of atoms possessing magnetic moments. This requires, however, that prior to the formation of the solid the atoms have not been linked to form diamagnetic molecules. But we shall discuss this later.

---

* In this case the magnetic moment is of a classical nature: an electron orbiting the nuclei produces a current loop that behaves as a microscopic magnet. By virtue of quantum laws, the projection of the angular momentum equals an integer number of Planck's constants.

Exchange interaction

The magnetic properties of crystals depend essentially on whether a unit cell possesses a magnetic moment or not. If it does, a crystal is said to be *paramagnetic*, and if it does not, the crystal is *diamagnetic*. Both types of magnetic crystals are common in nature.

Of course, if the unit cell contains a single atom, the magnetism of the crystal is completely determined by whether the atom is paramagnetic or diamagnetic. Solid neon is diamagnetic, while iron or chromium crystals are paramagnetic. But if the crystal is composed of regularly arranged molecules, then the presence or absence of a magnetic moment in a unit cell is determined by whether the molecule is paramagnetic or diamagnetic.

Let us try to clarify this far from trivial situation. We want to find out how individual atoms form a more complicated structure, a molecule.

In Chapter 1 (p. 72) we discussed the nature of the forces of interaction between atomic particles. The following fundamental fact should be recalled: both the forces of attraction and repulsion act on atoms; these forces have different natures but both are of electrostatic origin, with very small magnetic interaction between particles (one of the roles played by 1/137). At a certain distance, of the same order as the atomic dimension, repulsion and attraction are equalized and the net force is zero. This is how molecules are formed ... . We find that some molecules (such as hydrogen molecules) possess zero magnetic moment, while others (such as oxygen molecules) possess nonzero magnetic moment. What factor

is decisive here? A correct but rather vague answer is: the final state has the lowest energy. In other words, a molecule is paramagnetic if the paramagnetic state has the lowest energy, and it is diamagnetic if the diamagnetic state has the lowest energy. This is no doubt true. But why does the energy of a molecule at all depend on its magnetic moment even when a molecule is outside external magnetic fields?

The following answer immediately comes to mind. Is not each electron a small magnet? The energy of interaction between magnets depends upon their relative orientation (like poles are repelled, and unlike poles are attracted). Hence, having constructed a molecule out of atoms, that is, knowing the spatial configuration of electron orbits, one has to calculate the configuration energy for parallel magnetic moments and compare it with the configuration energy for antiparallel magnetic moments (when there are two electrons in a molecule, as in the hydrogen molecule). Yes, this is all correct, but does not explain some experimental facts. The difference in energies of the paramagnetic and diamagnetic configurations is in fact of the order of the energy of electrostatic interaction between electrons that are an atomic distance apart. This energy is definitely independent of the orientations of magnetic needles, and the magnetic energy is lower by a factor of tens of thousands than the electrostatic energy. This explanation thus will not do.

A correct explanation is much more complicated. A consistent quantum mechanical treatment shows that the state of electrons is determined not only by the character of their motion in the electric field of the nuclei but also by the relative

position of spins, even if the direct magnetic interaction is ignored. This phenomenon was called the *exchange interaction*, and that part of the energy of electrons that depends on the relative orientation of spins is called the *exchange energy*. The exchange energy of the paramagnetic state in the hydrogen molecule is higher than that of its diamagnetic state. The hydrogen molecule is diamagnetic. In oxygen, however, the paramagnetic state has the advantage from the energy point of view (again because of the exchange energy).

The nature of the exchange interaction is very difficult to explain in popular science terms because it has no classical analogue. It arises because of the indistinguishability of particles and the Pauli exclusion principle, and thus owes itself solely to quantum properties of electrons. As an example, we consider two atoms. Each atom contains one electron. An analysis of the state of the electrons shows that we cannot say that one electron is in atom No. 1, while the other is in atom No. 2. It is as if both electrons swap places all the time, as if they hop from one atom to another (indistinguishability). These hops are connected to a definite energy (*exchange* energy) that depends essentially on the orientation of spins (parallel or antiparallel spins). Allowed states are selected according to the Pauli exclusion principle. In some cases the exchange energy is higher for parallel spins, while in other cases it is higher for antiparallel spins.

Exchange forces decrease very rapidly with distance: the larger the distance to the neighbouring atom the more difficult it is for an electron to "hop" across the gap. It is important to empha-

size once again that exchange forces are not related to the usual interaction of magnetic spin moments (we have already mentioned that it is very low). Let us point out an important property of the exchange energy. As this energy depends on the *relative* arrangement of electrons' spins, the direction of the magnetic moments of the molecule becomes immaterial, provided the molecule is not placed in a magnetic field.

We have made one more step towards understanding magnetic properties. Nature atoms and molecules with nonzero magnetic moments. The energy values that are responsible for the existence of magnetic moments are quite high. They are determined by the electrostatic interaction between charges.

Paramagnetics and ferromagnetics

We thus know the building blocks of which magnetic materials are composed. These are atoms or molecules with nonzero magnetic moments.

The physical basis of the coalescence of atoms and molecules into a solid (a crystal) is the same as that which makes atoms combine into molecules. We are not going to discuss it here, but let us consider whether the fact that the molecules we are speaking about are paramagnetic, that is, possess each its own magnetic moment, plays any part in this process of coalescence. Strictly speaking, it does make a contribution. The exchange energy is not low and, therefore, its contribution to the binding energy of a crystal is appreciable. Nevertheless, the exchange energy is substantially lower than the binding energy and, therefore, as a first step we can simply forget

about it and assume that magnetic moments do not interact (very soon we shall discard this assumption). Therefore, a paramagnetic material which is formed by the coalescence of paramagnetic molecules or atoms must be seen as an array of regularly spaced but freely rotating nonfixed magnetic needles. The directions of the needles are random. True, quantum mechanics forbids certain directions, but once a "given" direction is "allowed", the reverse direction is allowed as well. As a result, the net magnetic moment equals zero.

There is a very simple method of "calling magnetic moments to order". Let us place a paramagnetic material in a magnetic field. The energy of interaction of magnetic moment μ with magnetic field H is $-\mu H$. If the magnetic moment can be oriented with respect to the field in only two ways (the case of the electron), then the gain in energy (the difference between the energies of states parallel and antiparallel to the field) equals $2\mu H$. Thermal motion disrupts the alignment of magnetic moments along the field and prevents complete ordering. Quantity $2\mu H$ must first be compared with the temperature in energy units, that is, with $k_B T$. Since $\mu \approx 10^{-20}$ erg/G, we find that the equality $2\mu H = k_B T$ is possible either at extremely low temperatures or fantastically strong magnetic fields (if $T = 1$ K, then the magnetic field at which $2\mu H = k_B T$ is approximately 10 000 G). At ordinary* and high temperatures $2\mu H$ is much smaller

---

* The word "ordinary" must be used with caution. What should be considered ordinary? In many a laboratory, temperatures below 1 K and magnetic field strengths above 100 000 Oe are quite routine nowadays.

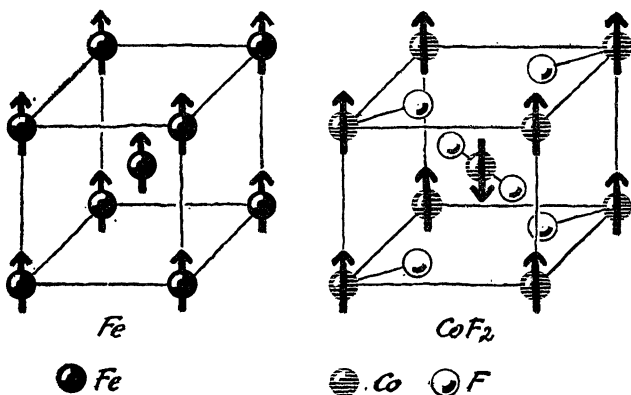Fe   CoF₂

● Fe   ◐ . Co   ◑ F

Fig. 37

than $k_B T$. As a result, a magnetic field only negligibly aligns magnetic moments.

However, there is an internal factor which aligns magnetic moments; this factor is the exchange interaction that was discussed in detail in the preceding section.

The exchange interaction aligns magnetic moments either in parallel or antiparallel to one another. In the first case we obtain a ferromagnetic, and in the second case, an antiferromagnetic.* This is illustrated schematically in Fig. 37.

If we compare the ordering effect of exchange forces with the effect of the magnetic field, we

---

* Numerous materials (they are called *ferrimagnetic*) can be found in nature which according to this classification occupy an intermediate position. Magnetic moments in such materials point in opposite directions. However, the number of "left-handed" and "right-handed" moments is not equal.

shall have to concede that atomic magnets are in a magnetic field of enormous strength—tens of millions of oersteds.

Let us for the present leave the discussion of antiferromagnetics and start with materials in which magnetic moments in a crystal interact in such a manner that it becomes advantageous for them, from the energy point of view, to be aligned in parallel to one another; obviously, the thermal motion constantly disrupts this alignment. As long as the temperature is high, the thermal motion is so intense that there is no ordering in the positions of the magnetic moments (the net magnetic moment is zero). However, at a certain temperature, the *Curie point*, denoted by Θ_C, the interaction becomes significant and the body gains a nonzero magnetic moment. Magnetic moments become aligned* spontaneously (not under the influence of the external magnetic field but because of internal factors). The lower the temperature the higher the spontaneous moment. The state in which an object is found at temperatures below the Curie point Θ_C is called the *ferromagnetic* state, and the object itself is called a *ferromagnetic*. At absolute zero all magnetic moments are aligned in the same direction (complete ordering). The magnetic moment per unit volume is then equal to the number of magnets per unit volume, N, times μ. Typical representatives of ferromagnetics are iron Fe (which gave the name "ferromagnetic" to the whole class), cobalt Co, and nickel Ni. Their spontaneous moments M_s at absolute zero are:

---

* The effect of the exchange interaction on the magnetic properties of materials was first pointed out by W. Heisenberg and J. Frenkel in 1928.

	M_s, erg/G	Θ_C, K
Fe	1735	1043
Co	1445	1403
Ni	509	631

Things we have ignored. Domains

In the real world everything works differently. If a paramagnetic is cooled outside an external magnetic field, its magnetization is zero even when its temperature drops below the Curie point Θ_C. Only an external field will magnetize it, and only by increasing the strength of this field can we reach total magnetization. The magnetization curve (magnetic moment of a body vs. magnetic field strength) is determined both by the shape of the specimen and by its state: whether it is very pure or contains impurities, whether it has been well annealed or has internal stresses.

A very important property of magnets, which is decisive for their application in industry, science, and household appliances, is their capability to retain magnetization after the external magnetic field has been switched off; owing to this property, magnets can serve as sources of magnetic field. It is as if a magnet accumulates magnetic energy which can be utilized in all sorts of ways.

But although all these problems are very important and very interesting they are very far from what concerns us here. The point of the

matter is thus. If a paramagnetic specimen were infinite,* the sequence of events would be exactly as that outlined in the preceding section: as the temperature drops below the Curie point, the paramagnetic spontaneously (sic!) gains a magnetic moment, which grows as the temperature decreases. However, owing to end effects things are more complicated in a finite specimen. The boundary of a magnetized specimen is a source (or sink) of magnetic lines of force. (Magnetic lines of force are never interrupted!) As a result, a magnet is always surrounded by a magnetic field that fades away rather slowly with distance. A magnetic field contains an energy. Hence, the energy of a ferromagnetic is somewhat higher than we thought before: it includes the energy of the magnetic field produced by the magnet.

Each body tends to minimize its energy as far as this is possible—it is possible for magnetic material. The exchange energy (the main cause of magnetization) is independent of the direction of the net magnetic moment. The following effect is then possible. The magnetic specimen "separates" into domains. The magnetic moment is uniform in each domain but (in the simplest case) the magnetic moments of neighbouring domains are oriented in the opposite directions.** On the whole, a body has zero magnetic moment and therefore produces no magnetic field. To be

---

* As a model of an infinite magnet we can take a torus large enough to completely enclose the magnetic lines of force.
** The concept of domains was first introduced by P. Weiss in 1907.

233

exact, the field exists but only at distances of the order of the domain size.

From the standpoint of the magnetic field, therefore, energy is minimized by dividing a magnetic specimen into the maximum possible number of domains. In fact, the existence of boundaries between domains is a disadvantage, since the magnetic moments on both sides of a domain-domain interface are directed against each other. This means that the exchange energy will be minimized by decreasing the number of boundaries and, hence, the number of domains. One energy is decreased by increasing the number of domains, while the other energy is increased. The sum of both the energies has a minimum for an intermediate number of domains. This is the number of domains into which a magnetic specimen is "separated". Calculations show (L. Landau and E. Lifshitz, 1935) that the domain size increases proportionally to the square root of the specimen dimensions.

Domains are no longer a hypothesis. They were well studied both theoretically and experimentally. There are many ways of visualizing domains (even to the naked eye!).

The appearance of domains is thus a secondary effect.

Internal forces magnetize the specimen and order the magnetic moments of atoms and molecules, while the tendency to diminish the energy of a system results in "separating" the specimen into domains. This is something that in the following we will ignore. We will either assume that we are dealing with an infinite magnetic specimen or that our object is a single domain. I repeat that the structure of magnetic materials is an

234

interesting and important problem. Neverthe less, our goal is to interpret the *microscopic* structure of magnetism ... .

Antiferromagnetics

In the thirties a group of young physicists in Kharkov, headed by L. V. Shubnikov, discovered striking anomalies in the behaviour of a number of substances (chlorides of transition elements). At a certain temperature their specific heat changed abruptly, while the magnetic susceptibility at the same temperature reached a maximum. Several years before L. Landau had theoretically analyzed the behaviour of a system of magnetic moments in which the exchange interaction is such that it is advantageous for them to be aligned in an antiparallel array. He had demonstrated that there is a threshold temperature Θ_N (later called the *Néel temperature**) below which the magnetic moments become, in a certain sense, ordered. The lower the temperature the higher the ordering. The arrangement reaches complete ordering at absolute zero. But this ordering is not that of a ferromagnetic. Each magnetic moment is surrounded by antiparallel neighbours (see Fig. 37). Therefore substances with this type of ordering were named *anti*ferromagnetics. According to Landau's theory, the anomalies that we have mentioned above should be observed, and were observed, close to the Néel temperature ... .

---

* Louis Néel was the French theoretical physicist who played an outstanding role in the study of antiferromagnetics. He was awarded the Nobel prize in physics for his contribution to magnetism.

Numerous antiferromagnetics are known in nature. Here are some examples taken from a work by A. S. Borovik-Romanov:

Compound	Θ_N, K	Compound	Θ_N, K
$NiSO_4$	37	FeO	188
$FeSO_4$	21	NiF_2	73.2
NiO	520	FeF_2	78.3

Not only compounds form antiferromagnetics. Chromium and many rare earth metals are antiferromagnetic as well.

Magnetic ordering in ferromagnetics is easily detected. The information is passed on by the magnetic field created by the ferromagnetic, or (if we drop the literary embellishments): the spontaneous magnetic moment of a ferromagnetic can be measured directly. And could we directly detect ordering in the arrangement of magnetic moments? This can be done by studying the scattering of neutrons by an antiferromagnetic.

Why neutrons, why not the routine X-rays or electrons which are applied so successfully to the structural analysis of crystals? An atom with its downward magnetic moment and an atom with its upward magnetic moment are practically indistinguishable both for X-rays and electrons. They only interact (better to say, mostly interact) with charges within the atom. The neutron has no charge, but its magnetic moment is nonzero. Therefore atoms with different direc-

tions of magnetic moments are not the same for neutrons, and differently oriented atoms scatter neutrons in a different manner.

Magnons in ferromagnetics

To find the arrangement of magnetic moments is to obtain only one half of the solution. For a full understanding of the properties of magnetics it is very important to know how magnetic moments move.

While studying the motion of magnetic moments, the motion of an individual magnetic moment cannot be considered, since this motion can be treated only collectively (we shall illustrate this statement with a specific example).

Let us consider the properties of magnetics in the vicinity of 0 K. Motion stops almost completely. It is the simplest form of motion. In particular, we can ignore changes in the magnetic moments of atoms, since the changes would be accompanied by a considerable increase in the energy of electrons within the atom; at low temperatures this is extremely unlikely. So the somewhat vague expression "the simplest form of motion" can be elaborated: the magnetic moments of atoms can "rotate" around fixed centres, but that is all. In fact, we have to put "rotate" in quotation marks: it is difficult to talk about rotation if the moment can assume only two positions in space. We shall begin with precisely this case, by choosing for the sake of simplicity a ferromagnetic consisting of spin 1/2 particles.

We shall consider weakly excited states of a ferromagnetic in which all magnetic moments in the lowest, ground state (0 K) are oriented

in the same direction. Excited states of a ferromagnetic are easily classified by the value of the deviation of the magnetic system from the completely ordered arrangement. Let us conduct a Gedanken experiment. We reverse the magnetic moment of one atom and "leave the ferromagnetic to its tricks". The exchange interaction makes the state with a flipped (reversed) magnetic moment disadvantageous: its magnetic neighbours will try to "call their colleague to order". However, and this is possibly one of the most important features, in so far as only exchange forces are taken into account, the ferromagnetic cannot change its net magnetic moment.* Therefore the tendency to call the neighbour to order results in flipping another adjoining magnetic moment. A wave of flipping magnetic moments, or spins, thus travels through the crystal. These waves, which were predicted by Bloch in 1930, are called *spin waves*.

An analysis of the propagation of waves of spin flips shows that the spin wave frequency ω and its wavelength $\lambda = 2\pi/k$ are related by a dispersion law that expresses frequency as a function of wave vector $\mathbf{k}$. Hence, a state with a single flipped magnetic moment is a wave with a certain wave vector $\mathbf{k}$ and a corresponding frequency $\omega(\mathbf{k})$.

The reader is already familiar with our next step: we use the de Broglie relations and pass from a wave to a quasiparticle. This quasiparticle is called *magnon*. A crystal with a single flipped spin is thus crystal with a single magnon.

---

* Rigorously formulated, the last statement must be: the system's spin is an integral of motion if only electrostatic forces operate.

Note that in mentioning a single flipped magnetic moment we cannot specify the position of this maverick moment. A spin wave "involves" all magnetic moments of the crystal.

The dispersion law is very simple for long wavelengths, that is, for small wave vectors: $\omega \sim k^2$.

The proportionality factor is readily estimated if we operate not with the frequency of a spin wave but with the magnon's energy $\varepsilon = I\,(ak)^2$, where a is the interatomic spacing. This formula shows that constant I coincides (of course, by order of magnitude) with the energy of the exchange interaction of neighbouring magnetic moments.* Numerically this energy lies close to the Curie point of the ferromagnetic expressed in energy units ($I \approx k_B \Theta_C$). We thus have a reliable criterion for the estimation of the most important variable in the dispersion law. True, this is a very approximate criterion: I may differ from $k_B \Theta_C$ by a factor of 2 to 3.

The formula for the magnon's energy is conveniently rewritten in a different form, if wave vector k is replaced by momentum p and a new constant with the dimension of mass is introduced:

$$\varepsilon = \frac{p^2}{2m^*}, \qquad m^* = \frac{\hbar^2}{2a^2 I}. \qquad (79)$$

The new constant, m^*, is called the *effective mass*. This last expression shows especially clearly that the magnon behaves very much like a particle.

---

* If $k \approx 1/a$, then the spins of neighbouring atoms are antiparallel.

Flipping two or more magnetic moments in a ferromagnetic crystal means propagating several spin waves, instead of a single wave. In corpuscular terms, several magnons exist in a ferromagnetic crystal. When the temperature increases, the number of magnons increases as well, while the net magnetic moment of the ferromagnetic crystal correspondingly diminishes. The interaction between magnons can be ignored as long as the number of spin waves (magnons) is small; these magnons can be treated as an ideal gas. The number of particles in this gas is the greater the higher the temperature. The techniques of quantum statistical physics enable us to study the properties of this gas and establish that the magnon gas is governed by the Bose-Einstein statistics.

Magnons are bosons. Their chemical potential naturally equals zero. Knowledge of their statistical properties suggests a method of calculating their macroscopic characteristics. It was thus established that the number of magnons is proportional to $T^{3/2}$. This means that the spontaneous magnetic moment of a ferromagnetic, M, diminishes at low temperatures proportionally to $T^{3/2}$; indeed, each magnon is nothing but one flipped magnetic moment travelling in the ferromagnetic:

$$M(T) = M_0 \left[1 - \beta \left(\frac{T}{\Theta_C} \right)^{3/2} \right],$$

$$T \ll \Theta_C, \quad \beta \approx 1, \quad M_0 = N\mu. \tag{80}$$

The magnon gas contributes to the thermal properties of ferromagnetic materials. At low temperatures ($T \ll \Theta_C$) the specific heat contains, owing to magnons, a term also proportional to

$T^{3/2}$, that is, again proportional to the mean number of the magnons. The linear dependence of the thermodynamic properties on temperature to the power 3/2 is often called the "law of 3/2". It must be emphasized that the contribution of magnons to the specific heat is larger at low temperatures than that of phonons, because the specific heat of the phonon gas decreases with temperature faster, viz. proportionally to T^3.

Both phonons and magnons are bosons, so both gases have zero chemical potential, $\zeta = 0$. The difference in the thermal behaviour of specific heat is a result of the difference in dispersion laws of phonons and magnons ($\varepsilon = sp$ for phonons and $\varepsilon = p^2/2m^*$ for magnons).

Magnons are "full-blooded" quasiparticles. They transfer heat, participate in the absorption of sound wave energy, and act as obstacles scattering electrons in ferromagnetic metals. These properties of magnons are far from hypothetical. Experiments reveal the contribution of magnons to various properties of ferromagnetic crystals. In particular, the temperature dependence of the magnetic moment and the contribution of magnons to specific heat have been reliably confirmed by experiments. The contribution of magnons to the electron scattering in metal ferromagnetics has also been identified. We are able to observe a single magnon. But this will be taken up later.

Magnons in antiferromagnetics

At $T = 0$ K all magnetic moments of atoms in an antiferromagnetic are also strictly ordered, only the ordering is more complicated. Flipping any of the magnetic moments is a perturbation

which propagates in a crystal as a wave. However, the existence of two types of atoms (first, with spins "up", second, with spins "down") produces two types of spin waves which can propagate in antiferromagnetic crystals; that is, two types of magnons can appear. As long as only the exchange interaction is taken into account, the properties of these magnons are almost indistinguishable: they have dispersion laws essentially different from those of magnons in ferromagnetics. Magnons in antiferromagnetics behave very similarly to phonons. For small values of momentum $\mathbf{p} = \hbar\mathbf{k}$,

$$\varepsilon_{1,2} \approx \Theta_N ak, \quad ak \ll 1. \tag{81}$$

Subscripts 1 and 2 emphasize that there are two types of magnons. We recall that Θ_N is the Néel temperature, and a is the interatomic spacing. Magnons in antiferromagnetics are again bosons.

As the temperature increases, the number of magnons increases as well ($\zeta = 0$), thereby gradually disrupting the rigorous ordering in the alignment of magnetic moments. The increase in the number of magnons with temperature leaves its imprint on the thermal and magnetic properties of antiferromagnetics.

The difference in the dispersion laws for magnons in ferro- and antiferromagnetics is responsible for the difference in temperature dependences. Thus, the magnetic part of the specific heat of an antiferromagnetic is proportional to T^3 as in the case of phonons, not to $T^{3/2}$ as in the case of ferromagnetics.

An analysis of the properties of magnons in antiferromagnetics demonstrates that they are

very sensitive to the structure of the ground state. A change in the arrangement of the magnetic moments produces alterations in the dispersion law of magnons. The most natural influence of magnetic moments on the structure is produced by the magnetic field, which "attempts" to align magnetic moments along its own direction. Therefore the investigation of the magnetic properties of magnetic materials in sufficiently strong magnetic fields is one way of studying the properties of magnons.

Magnons from a slightly different standpoint

Now that we understood the microscopic nature of spin waves as waves of spin flips, we can try to describe them in simpler terms, by using macroscopic concepts.

We consider the uniform rotation of magnetic moments, that is, an infinitely long spin wave. This means that all atoms are subject to absolutely identical conditions, in other words, exchange forces are, so to speak, switched off. If no other forces were applied to the magnetic moments in a body, the rotation of magnetic moments would not absorb energy. This means (the de Broglie relation comes into operation again) that the frequency of the uniform rotation of magnetic moments is zero when only exchange forces are considered. But in addition to exchange forces, there are other forces which are not indifferent to the direction of magnetic moments. They are called anisotropy forces (it is more customary to refer to the *anisotropy energy*). A good idea of the anisotropy force can be gained

from an imaginary magnetic field H_a applied along the preferred axis of the ferromagnetic crystal. Such an axis is always directed along one of the crystallographic axes of the ferromagnetic crystal. The spontaneous magnetic moment is always directed along this axis. In the case of cobalt this is the six-fold axis of the hexagonal lattice, and in the case of iron these are three axes along the edges of the cubic unit cell.

What will happen to the magnetic moment in the magnetic field if we tilt it by some angle? The magnetic moment will rotate around the magnetic field because of the moment of force applied to it perpendicularly to the plane passing through the magnetic field vector and the magnetic moment itself. The frequency of rotation ω around the field H_a is equal to $\omega = gH_a$ for spin 1/2 particles; here $g = e/mc$ is the quantum gyromagnetic ratio (see p. 60).

The actual magnetic field H applied along the preferred axis increases the frequency of rotation because it is added to the anisotropy field H_a, that is, $\omega = g\,(H_a + H)$.

If the deflection of magnetic moments is not uniform (say, if we induce the rotation of the magnetic moments located along the boundary of the specimen), exchange forces will be switched on and this will trigger the propagation of a wave along the ferromagnetic; the frequency of this wave is the sum of the frequencies of the uniform and nonuniform rotation:

$$\omega = g\,(H_a + H) + \frac{I}{\hbar}\,(ak)^2. \qquad (82)$$

The existence of the anisotropy energy or of the energy of the external magnetic field leads to

244

nonzero frequency of the spin wave at $k = 0$. Or, in quantum terms, a magnon with zero momentum has nonzero energy, viz. the "rest energy" equal to

$$\varepsilon_0 = \hbar g (H_a + H) = 2\mu (H_a + H).$$

The rest energy of ordinary particles is mc^2, where m is the particle's mass, and c is the speed of light. Thus the calculated rest mass of a magnon is $m_0 = \varepsilon_0/c^2$. This quantity is fantastically small: $m_0 \approx 10^{-37}$ g at $H = 10\,000$ Oe (this H, as a rule, exceeds H_a), that is, the magnon is by a factor of 10^{10} smaller than the lightest of the particles, the electron. Nevertheless, the rest energy of magnons is detected in some experiments. First, there are experiments in which the high-frequency properties of ferromagnetics are studied (they will be discussed later) and, second, experiments studying thermal and magnetic properties at very low temperatures.

If a ferromagnetic is cooled to extremely low temperatures, the thermal energy is insufficient for the creation of magnons. As a result, the number of magnons is exponentially small (proportional to exp $(-\varepsilon_0/k_B T)$), and the temperature dependencies of the spontaneous moment and the magnon component of the specific heat are modified: in the vicinity of absolute zero the "law of 3/2" is replaced by the exponential dependence on temperature. The exponent includes the rest mass of the magnon. It is possible to "weigh" a particle whose mass is smaller than that of the electron by a factor of 10^{10}!

Let us now turn to antiferromagnetics. If it is advantageous for magnetic moments to align along the axis of anisotropy (Fig. 38), then they
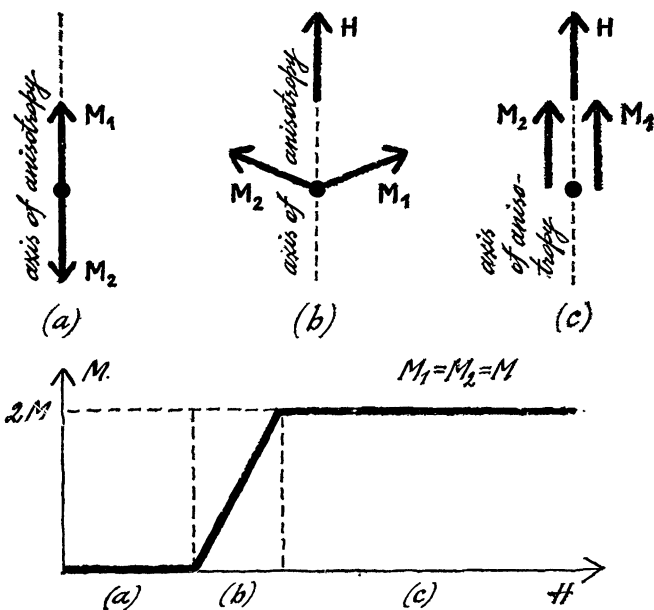
Fig. 38

will begin to rotate in the case of any deflection from the axis. There are two moments, so there are two frequencies:

$$\omega_1 = g\left(\sqrt{\overline{H_a H_{ex}}} + H\right),$$

$$\omega_2 = g\left(\sqrt{\overline{H_a H_{ex}}} - H\right),$$

where H_{ex} is the "exchange" field introduced in order to standardize notations: $\mu H_{ex} = I$.

Each of these frequencies (ω_1 and ω_2) corresponds to a comparatively complicated rotation of both moments. Both frequencies coincide when

the magnetic field is zero and thus are equal to $\omega_{1,2} = g\sqrt{H_a H_{ex}}$. The anisotropy field H_a is approximately the same as in ferromagnetic crystals, but the exchange field H_{ex} is much stronger than the anisotropy field. Therefore at $H \ll \sqrt{H_a H_{ex}}$ the frequencies of spin waves in antiferromagnetics are much higher than the corresponding frequencies in ferromagnetics. The external magnetic field raises one frequency and diminishes the other. It is an ally of the anisotropy energy for one of the rotations, and an adversary for the other. The last formulas are valid only if $H < \sqrt{H_a H_{ex}}$. If $H \approx \sqrt{H_a H_{ex}}$, the magnetic system in Fig. 38 becomes unstable and undergoes restructuring. Figure 38b shows what happens after restructuring. A further increase in the magnetic field induces a better alignment of the magnetic moments along the axis. And, finally, the magnetic moments flip over to the axis at a magnetic field equal to H_{ex} (Fig. 38c). The strength of the threshold field is in fact very high: energy μH is approximately equal to the exchange energy $I \approx k_B \Theta_N$. This means that the flip-over can be observed only in antiferromagnetics for which the Néel temperatures are low. The restructuring is clearly seen in the dependence of the net magnetic moment of the antiferromagnetic crystal (magnetization) on the magnetic field (the corresponding curves are given in the same figures).

The exponential dependence of the magnon components of thermodynamical variables is observed in antiferromagnetics at high temperatures because of the comparatively high frequency of uniform rotation. The rest mass of magnons
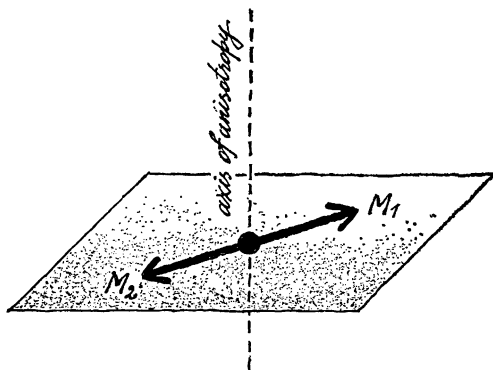
Fig. 39

in antiferromagnetics exceeds that in ferromagnetics.

This is true only for those antiferromagnetics in which it is advantageous for magnetic moments to align along the axis. There are other antiferromagnetics (we have discussed only a few so far) in which it is advantageous for magnetic moments to align in the plane perpendicular to the axis (Fig. 39). It is immaterial how the magnetic moments are oriented in the plane. In other words, the rotation of a system of magnetic moments around the axis of anisotropy does not consume energy. Consequently, one of the frequencies of rotation of magnetic moments is zero. The other frequency is approximately the same as in antiferromagnetics in which magnetic moments are parallel to the axis.

We have described only frequencies of uniform rotation. If the frequency of uniform rotation is

multiplied by Planck's constant, we obtain the rest energy of a magnon. Nonuniform rotations (with a wave vector distinct from zero, or a finite wavelength) correspond to travelling magnons, but enough has been said about them above.

We recapitulate the contents of this section: magnons with very low momentum are rotation waves of magnetic moments; magnons with zero momentum represent a uniform (in-phase) rotation of all magnetic moments of ferro- or antiferromagnetics.

Resonances

Magnetic moments of a ferromagnetic can be made to rotate by external means. In order to achieve this, we place a ferromagnetic in a rotating magnetic field. If the frequency of the rotation of the magnetic field differs from the natural frequency of the rotation of the magnetic moments, the magnetic field will alternately accelerate the magnetic moments (transfer energy to them) and decelerate them (absorb energy from them). On the average, magnetic moments will not absorb any energy from the rotating magnetic field, and an experiment carried out specially to measure the energy of the alternating magnetic field absorbed by a ferromagnetic will yield no result. But if the frequency of the alternating magnetic field coincides with that of the rotation of the magnetic moments, *resonance* will set in and the energy of the magnetic field will be absorbed.

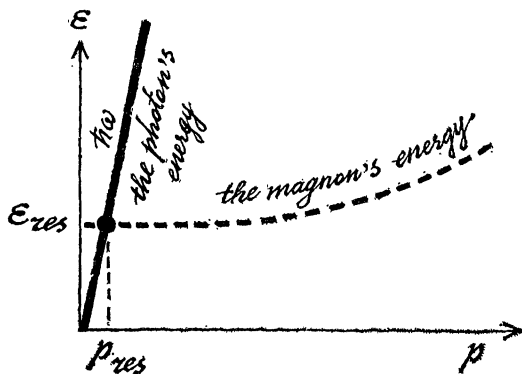This resonance is called *ferromagnetic resonance*. It was discovered by J. Griffiths in 1946.

Fig. 40

The frequency at which ferromagnetic resonance is observed depends essentially on the anisotropy energy and on the applied magnetic field. In order to clarify the orders of magnitude of the quantities involved we note that a magnetic field $H = 10\,000$ Oe corresponds, via the formula $\omega = gH$, to a frequency $\omega \approx 10^{11}$ s^{-1}, or a wavelength $\lambda \approx 6$ mm.

Ferromagnetic resonance has a natural quantum mechanical interpretation: a photon with momentum $p = \hbar\omega/c$ (where ω is the frequency of the electromagnetic wave) is transformed into a magnon whose energy ε is a function of p.

The transition can occur if the energy and momentum conservation laws are satisfied. Whether this is so is readily checked by means of Fig. 40 in which the energies of a photon and a magnon are plotted as functions of momentum. The curves intersect at that value of momentum at which a photon can be resonantly transformed

250

into a magnon. But the following has to be taken into account: we had to distort the graph because the speed of light is so high that the straight line showing the photon's energy as a function of momentum must virtually coincide with the y-axis. And this means that the magnon created in the photon-magnon transition has nearly zero momentum. In other words, ferromagnetic resonance makes it possible to measure directly the rest energy of magnons ("weighing" of the magnon's rest mass).

We have deliberately simplified our story. In reality the information yielded by experiments on ferromagnetic resonance is not limited to finding the frequency of uniform rotation. The energy absorbed by the ferromagnetic differs from zero not only at exact resonance but also in the vicinity of the resonance value of the magnetic field. The cause of this effect is as follows. The rotation of a magnetic moment cannot last indefinitely (in quantum terms, the lifetime of a magnon is not infinitely long). There are always processes which result in slowing down the rotation of magnetic moments. We did not mention this in the preceding section because the slowing-down time τ is, as a rule, much longer than the rotation period of the moment. According to the uncertainty principle, the lifetime in a given state, τ, and uncertainty in the energy of this state, $\Delta \varepsilon$, are related by the formula

$$\Delta \varepsilon \geqslant \frac{\hbar}{\tau}. \tag{83}$$

This means that the magnon's energy is known within $\Delta \varepsilon = \hbar/\tau$. Therefore, the width of the curve $\varepsilon = \varepsilon(p)$ in Fig. 40 is $\Delta \varepsilon = \hbar/\tau$. Obvious-

ly, the uncertainty in the magnon's energy makes possible "photon-magnon" transitions in a certain range of parameters. The second corollary is that, by studying the shape of a resonance curve (absorbed power as a function of magnetic field), we can extract information on processes which limit the magnon's lifetime. This is extremely interesting because it is indicative of the magnon's interaction with the members of its own family and quasiparticles of other "breeds", such as phonons. Unfortunately, we cannot dwell on this phenomenon here.

Ferromagnetic resonance is a method of finding the magnon's rest energy. But is a direct study of the properties of an *individual* mobile magnon possible? We emphasize that we mean an individual magnon. An ensemble of magnons is available for experimental study (this has been mentioned above), but conclusions always gain in "conclusiveness" if the object of study can be "isolated" and subjected to a "personal" scrutiny. Nature indeed has given us this opportunity.

Let us consider some of these possibilities.

We first consider magnons that are nearly at rest—magnons with very low momenta; in other words, we consider spin waves with very long wavelengths.

We recall that because of the tremendous velocity of a photon, it is turned into a zero-momentum magnon, not into a moving magnon (see Fig. 40). And what if a quantum of sound, a phonon, is made to transform into a magnon? The speed of sound is one hundred thousand times smaller than that of light. An analysis of the conservation laws shows that indeed *ferroacoustic resonance* (this is the name given to the transi-

tion of a phonon into a magnon) creates a magnon with a finite momentum. Ferroacoustic resonance was predicted by A. I. Akhiezer, V. G. Bar'yakhtar, and S. V. Peletminsky in 1956 and became one of the most effective methods of studying magnetics. Ferroacoustic resonance suggests one additional possibility for studying magnons with finite momenta. In order to better grasp the idea of the method, let us describe the creation of a magnon by a phonon in "wave terms": a sound wave has been transformed into a spin wave of the same wavelength. On the atomic scale, the wavelength of radio waves, even ultrashort waves used in ferromagnetic resonance experiments, is very long. Therefore in ferromagnetic resonance an infinitely long spin wave is excited (we should rather say that the wavelength of the spin wave is so long that it can be considered infinite). But can an electromagnetic wave be compressed, that is, shortened? A medium with a refractive index $\mathfrak{N}$ diminishes the light wavelength by $\mathfrak{N}$ when compared with vacuum propagation of light. But $\mathfrak{N}$ cannot be greater than a few units. Both light and radio waves are electromagnetic oscillations. Therefore we cannot really hope to contract the spin wavelength˚ on the basis of the refractive index. Metals, however, are an exception. They are almost impenetrable by electromagnetic waves, and that tiny fraction that does penetrate into the metal is damped out over very short distances. This distance (the so-called depth of the skin layer) is of the order of 10^{-5} cm for wavelengths of the order of 1 cm, that is, the electromagnetic field in a metal undergoes considerable changes over distances 100 000 times shorter

than the wavelength of the field in a vacuum. If we wanted to represent the electromagnetic field as an ensemble of plane waves, we would find that the main contribution is made by waves with a wavelength of $\approx 10^{-5}$ cm, while the wave incident on the metal surface has a wavelength of 1 cm. In other words, the metal contracts electromagnetic waves enormously, although it does also "blur" them: instead of a single plane wave we obtain a packet, that is, a system of waves whose superposition gives a field that is damped out in the bulk of the metal.

The contraction of the electromagnetic wave in metals makes it possible to use the study of high-frequency properties in metal ferromagnetics in order to analyze mobile magnons. To achieve this, it is necessary to measure the reflection coefficient of the electromagnetic wave as a function of frequency and the external magnetic field. Of course, the measurements must be made under conditions close to those of ferromagnetic resonance. The fact that an electromagnetic wave generates a spin wave and not the uniform rotation of magnetic moments affects the value of the reflection coefficient and enables us to find the properties of mobile magnons. Such experiments were first carried out by the American physicists W. Ament and G. Rado in 1955.

We have not mentioned so far a method which looks like the most natural way of studying finite wavelength spin waves: the resonance excitation of standing spin waves. Indeed, a ferromagnetic specimen is a resonator cavity for spin waves; reflection from the walls of this cavity must result in formation of standing waves. The thinner the film the more pronounced the inter-

ference pattern. Suitable ferromagnetic films must be about 10^{-6}-10^{-5} cm thick. Of course, it is not easy to prepare high-quality films with a thickness of several hundred atomic layers. Nevertheless, this experiment has been carried out. The experimenters studied the resonance absorption of the energy of an alternating magnetic field in a thin ferromagnetic film. Instead of a single peak corresponding to the excitation of uniform rotation of magnetic moments, they found several peaks; the interval between the peaks makes it possible to measure a quantity directly related to the exchange interaction between neighbouring atoms.

We thus find that several techniques exist for the investigation of long wavelength spin waves. Each of them has its advantages. All of them supply physicists with numerous data which pave the road to deeper penetration into the nature of the magnetic state. All of these methods share a common shortcoming. It is impossible with these methods (at the very least, it is extremely difficult) to analyze spin waves with wavelengths of the order of the interatomic spacing (in the case of such spin waves it is more correct not to drop the quasiparticle term, the magnons). But the thermal and magnetic properties of magnetic crystals are determined by all the magnons. Our picture would be incomplete if we limited it to nearly fixed magnons. One way to study magnons with high momenta is to observe the inelastic scattering of neutrons.

In principle, the dispersion law for magnons is reconstructed from neutron-magnon scattering in the same way as in the case of phonon scattering. This was described in Chapter 2.

We know that the uniform rotation of magnetic moments is possible not only in the ferromagnetics we have been discussing in this section but in antiferromagnetics as well. Hence, the resonance absorption of the energy of an alternating magnetic field is equally possible. This statement sounds particularly conclusive if formulated in quantum terms: a photon interacting with an antiferromagnetic may be transformed into a magnon. In essence, the physical nature of antiferromagnetic resonance is identical to that of ferromagnetic resonance. Therefore there is no need to repeat all that has been said before. However, it is interesting to underline the differences caused by the antiferromagnetic ordering of magnetic moments.

The spectrum of an antiferromagnetic is more complicated than that of a ferromagnetic. The rest energy of a magnon in most antiferromagnetics is much higher than that in a ferromagnetic. In the case of an antiferromagnetic, the resonance is either in an unfeasible or barely feasible range of frequencies, or one has to employ extremely strong magnetic fields, since one of the frequencies of the nonuniform rotation decreases as the external magnetic field increases. And, finally, the magnetic structure of antiferromagnetics is restructured by the magnetic field (this has been mentioned already); this restructures the magnetic's spectrum and therefore is manifested in antiferromagnetic resonance.

Antiferromagnetic resonance was first observed by a group of Dutch physicists in 1951 and later became a widely used method of studying antiferromagnetism.

256

* * *

We began this chapter by discussing the nature of magnetism and the logic of the exposition has led us to describing the high-frequency properties of magnetics. This is not accidental. A study of high-frequency properties is one of the most effective techniques of determining the energy structure.

We had to restrict the exposition to the simplest, most lucid phenomena and properties of magnetics. Bulky monographs are devoted to the high-frequency behaviour of magnetics; dozens of articles dealing solely with this topic are published literally every day, and international conferences and specialized workshops are regularly organized. Many discoveries have found their way into industry and are now reliably employed in radio, radar, computers, and in various relay devices. There is also constant progress in our understanding of the nature of magnetism in various solids. Extremely interesting features are being discovered. For example, it has been found recently that the structure of the ground state of a large number of magnetics is more complicated than it had been generally assumed. There are crystals in which magnetic moments tend to align in an almost antiparallel manner. They are nearly antiferromagnetic. They possess magnetic moment in the ground state (even in the absence of external magnetic fields), but this moment is very low, many times smaller than in typical ferromagnetics. This phenomenon, discovered by A. Borovik-Romanov, was called *weak ferromagnetism* and was explained by I. Dzyaloshinsky (1957). In other

257

crystals the structure of the ground state is especially bizarre: the magnetic moments of neighbouring unit cells are rotated a little with respect to one another, so that a helix is formed with a pitch not equal to an integer number of unit cell dimensions. It has been found that conduction electrons play an important role in the formation of the magnetic helix in metals.

Our profound understanding of the nature of magnetism does not mean at all that each problem has been solved. Someone has compared the progress of science with a military offensive. As a rule, pockets of resistance are left behind the advancing lines. However, in contrast to military operations, we never know whether the pocket of resistance left behind will soon become a new frontline. Science has nothing comparable with military intelligence which could make such predictions possible. A new discovery, a new theory may force a drastic redistribution of our forces.

Concluding Remarks

You cannot become familiar with a country if you only look at it through the bus window, stay in wayside hotels, and do tourist sightseeing "on the run". However, the tourist trip may be all you can get and you know there will be no other chance. You travel across several regions of an enormously large country. The guide—the author—will consider his mission fulfilled if you felt that the way of life you saw did not irritate you, if you liked the places you saw, and if you were attracted to those only mentioned in the guidebook.

A popular science book always resembles an advertisement: tedious details are hidden and difficulties are mentioned only in passing. I hope that the reader realizes that this story of the quantum physics of the solid state presents no more than the façade of this science, that the methods of theoretical physics are not limited to more or less conclusive analogies and estimates. Theoretical physics is a rigorous science employing subtle physical arguments and painstaking, sometimes extremely difficult, mathematical cal-

culations. Only if a scientist is at home with the techniques and concepts of modern physical theory can he hope to obtain new results. Moreover, the results are a fruit of a lot of work, as in any other field of human activity; fortunately, such work brings joy not only from the results but also in the process of trying to solve problems that have not yet been solved.

TO THE READER

Further suggested popular science books
on physics from Mir Publishers
to be issued in 1981

LASER AGE IN OPTICS

L. TARASOV, Cand. Sc.

This is a popular science book concerned with new trends
in optics arising from the development of the laser. The
author successfully combines a popular style of presenta-
tion with, when necessary, a serious scientific treatment
of the problems involved. Consequently, the book can be
recommended not only for high school students but also
for a layman adult reader wanting to study in some
detail the principles involved in lasers, optical holography
and nonlinear optics. The reader will find a popular,
vivid and intelligent presentation of many elusive con-
cepts which, as a rule, are either not covered at all in
literature of this kind or only in passing.
The reader will close the book with the feeling that
he has just passed through an exciting field of physics in
which the three main areas of research form an inter-
twined and mutually enriching complex, bringing new
wonders to the already magical world of science and
technology.
Contents. From Incoherent to Coherent Optics. Optical
Holography. Nonlinear Optics.

PHYSICS FOR EVERYONE (in 4 books)
Book 3. Electrons

A. KITAIGORODSKY, D. Sc.

The fourth Russian edition of *Physics for Everyone* by
Lev Landau and Alexander Kitaigorodsky was pub-
lished in 1978 as two separate books: *Physical Bodies*
(Book 1) and *Molecules* (Book 2). They were pub-
lished in English in 1980 period. This is the first publica-
tion of Book 3 of this series. It was written as a sequel
to *Physics for Everyone*. This book deals with physical
phenomena in which our attention is focused on the
next level in the structure of metal—the electrical struc-
ture of atoms and molecules. Electrical and radio en-
gineering, without which the existence of today's civ-
ilization is inconceivable, are based on laws govern-
ing the motion and interaction of electrical particles,
primarily electrons, the quanta of electricity.
The main subjects of this book are electric currents,
magnetism and electromagnetic fields.
Contents. Electricity. Electrical Structure of Matter.
Electromagnetism. Summary of Electrical Engineering.
Electromagnetic Fields. Radio.

Book 4. PHOTONS AND NUCLEI

A. KITAIGORODSKY, D. Sc.

The book concludes the series *Physics for Everyone* by
the world-renowned scientist, winner of the Nobel and
Lenin prizes, academician Lev Landau and the distin-
guished physicist Alexander Kitaigorodsky.
This book discusses in a simple and easy-to-understand
manner the phenomena of electromagnetic waves, thermal
radiation, and current treatment of spectroscopic anal-
ysis. Provides an introduction to the field of nuclear
physics and explains the most common types of lasers.
Outlines principal aspects of special theory of relativity
and quantum mechanics.
Contents. Soft Electromagnetic Radiation. Optical Instru-
ments. Penetrating Electromagnetic Radiation. The
Unification of Mechanics. Structure of Atomic Nuclei.
Energy Surrounding Us. Physics of the Universe.

The book introduces the reader to the fascinating
world of solid state physics as governed by quantum
mechanics. It gives a clear exposition of the dynamics
of the crystal lattice (propagation of waves, heat conduc-
tion, and similar problems), of electric phenomena in
crystals in conjunction with the properties of the lattice,
and of the magnetic properties of crystals.
The author, using almost no formulas, manages to
give a complete explanation of the nature of pheno-
mena. An attentive reader will find in this book a lot
more than is found on the surface and will appreci-
ate its very personal style.